DOCUMENT RESUME

ED 070 657	24	SE 015 466

AUTHOR	Leffin, Walter William
TITLE	A Study of Three Concepts of Probability Possessed by Children in the Fourth, Fifth, Sixth and Seventh Grades.
INSTITUTION	Wisconsin Univ., Madison. Research and Development Center for Cognitive Learning.
SPONS AGENCY	National Center for Educational Research and Development (DHEW/OE), Washington, D.C.
REPORT NO	WURDCCL-TR-170
BUREAU NO	BR-5-0216
PUB DATE	Sep 71
CONTRACT	OEC-5-10-154
NOTE	281p.
EDRS PRICE	MF-$0.65 HC-$9.87
DESCRIPTORS	Curriculum; *Doctoral Theses; *Elementary School Mathematics; *Mathematics Education; *Probability; *Research

ABSTRACT

The status of three basic concepts of probability--points of a finite sample space, probability of a simple event in a finite sample space, and quantification of probabilities--possessed by children in grades four through seven was examined. A test for each of the three concepts was constructed by the author and administered to a total of 528 students who had not been taught probability previously. A multivariate analysis of covariance was performed on the results of the three tests, with grade equivalent scores on the Stanford Arithmetic Achievement Test used as covariates. The overall mean performances were significantly different among I.Q. groups, sex groups, and grades. The children demonstrated that they had acquired considerable knowledge about the three probability concepts without having received formal training. (Author/DT)

TECHNICAL REPORT

THE WISCONSIN RESEARCH
AND DEVELOPMENT CENTER
FOR COGNITIVE LEARNING

The University of Wisconsin Madison, Wisconsin

U S. Office of Education Center No. C-03 Contract OE 5-10-154

Technical Report No. 170

A STUDY OF THREE CONCEPTS OF PROBABILITY POSSESSED BY CHILDREN IN THE FOURTH, FIFTH, SIXTH AND SEVENTH GRADES

Report from the Project on Prototypic Instructional Systems in Elementary Mathematics

By Walter William Leffin

H. Van Engen, Principal Investigator

Wisconsin Research and Development
Center for Cognitive Learning
The University of Wisconsin
Madison, Wisconsin

September, 1971

This Technical Report is a doctoral dissertation reporting research supported by the Wisconsin Research and Development Center for Cognitive Learning. Since it has been approved by a University Examining Committee, it has not been reviewed by the Center. It is published by the Center as a record of some of the Center's activities and as a service to the student. The bound original is in The University of Wisconsin Memorial Library.

Published by the Wisconsin Research and Development Center for Cognitive Learning, supported in part as a research and development center by funds from the United States Office of Education, Department of Health, Education, and Welfare. The opinions expressed herein do not necessarily reflect the position or policy of the Office of Education and no official endorsement by the Office of Education should be inferred.

Center No. C-03 / Contract OE 5-10-154

2

National Evaluation Committee

Helen Bain
Immediate Past President
National Education Association

Lyle E. Bourne, Jr.
Institute for the Study of Intellectual Behavior
Universiy of Colorado

Jeanne S. Chall
Graduate School of Education
Harvard University

Francis S. Chase
Department of Education
University of Chicago

George E. Dickson
College of Education
University of Toledo

Hugh J. Scott
Superintendent of Public Schools
District of Columbia

H. Craig Sipe
Department of Instruction
State University of New York

G. Wesley Sowards
Dean of Education
Florida International University

Benton J. Underwood
Department of Psychology
Northwestern University

Robert J. Wisner
Mathematics Department
New Mexico State University

Executive Committee

William R. Bush
Director of Program Planning and Management and Deputy Director, R & D Center

Herbert J. Klausmeier, Committee Chairman
Director, R & D Center

Wayne Otto
Principal Investigator
R & D Center

Robert G. Petzold
Professor of Music
University of Wisconsin

Richard A. Rossmiller
Professor of Educational Administration
University of Wisconsin

James E. Walter
Coordinator of Program Planning
R & D Center

Russell S. Way, ex officio
Program Administrator, Title III ESEA
Wisconsin Department of Public Instruction

Faculty of Principal Investigators

Vernon L. Allen
Professor of Psychology

Frank H. Farley
Associate Professor
Educational Psychology

Marvin J. Fruth
Associate Professor
Educational Administration

John G. Harvey
Associate Professor
Mathematics

Frank H. Hooper
Associate Professor
Child Development

Herbert J. Klausmeier
Center Director
V. A. C. Henmon Professor
Educational Psychology

Stephen J. Knezevich
Professor
Educational Administration

Joel R. Levin
Associate Professor
Educational Psychology

L. Joseph Lins
Professor
Institutional Studies

Wayne Otto
Professor
Curriculum and Instruction

Thomas A. Romberg
Associate Professor
Curriculum and Instruction

Peter A. Schreiber
Assistant Professor
English

Richard L. Venezky
Associate Professor
Computer Science

Alan M. Voelker
Assistant Professor
Curriculum and Instruction

Larry M. Wilder
Assistant Professor
Communication Arts

GPO 826-078

Statement of Focus

The Wisconsin Research and Development Center for Cognitive Learning focuses on contributing to a better understanding of cognitive learning by children and youth and to the improvement of related educational practices. The strategy for research and development is comprehensive. It includes basic research to generate new knowledge about the conditions and processes of learning and about the processes of instruction, and the subsequent development of research-based instructional materials, many of which are designed for use by teachers and others for use by students. These materials are tested and refined in school settings. Throughout these operations behavioral scientists, curriculum experts, academic scholars, and school people interact, insuring that the results of Center activities are based soundly on knowledge of subject matter and cognitive learning and that they are applied to the improvement of educational practice.

This Technical Report is from Phase 1 of the Project on Prototypic Instructional Systems in Elementary Mathematics in Program 2. General objectives of the Program are to establish rationale and strategy for developing instructional systems, to identify sequences of concepts and cognitive skills, to develop assessment procedures for those concepts and skills, to identify or develop instructional materials associated with the concepts and cognitive skills, and to generate new knowledge about instructional procedures. Contributing to the Program objectives, the Mathematics Project, Phase 1, is developing and testing a televised course in arithmetic for Grades 1-6 which provides not only a complete program of instruction for the pupils but also inservice training for teachers. Phase 2 has a long-term goal of providing an individually guided instructional program in elementary mathematics. Preliminary activities include identifying instructional objectives, student activities, teacher activities materials, and assessment procedures for integration into a total mathematics curriculum. The third phase focuses on the development of a computer system for managing individually guided instruction in mathematics and on a later extension of the system's applicability.

ACKNOWLEDGMENTS

Sincere appreciation is extended to Professor Henry Van Engen for the guidance and encouragement given to me during my years of graduate study and for his valuable advice and guidance concerning this research study.

Gratitude is expressed to Professor M. Vere DeVault and Professor J. Fred Weaver who served on my committee.

Gratitude is also extended to Dr. Thomas Romberg, Dr. Margaret Harris, Mary Quilling and James Bavry for their helpful advice and interest in this study.

Further, my thanks are extended to the administrators, teachers and children of the Wausau Public School System where this study was conducted.

My deepest appreciation is expressed to my wife, Joyce, and our three children whose patient understanding and encouragement helped make the completion of this study possible.

TABLE OF CONTENTS

CHAPTER		PAGE
I.	BACKGROUND OF THE PROBLEM	1
	Introduction	1
	Summary of Recommendations for Teaching Probability in the Secondary and Elementary Schools	3
	Need for Research	10
	Interpretations of the Term "Probability"	13
	Interpretation of Probability to be Used in this Study	15
	Related Research	16
II.	THE PROBLEM	25
	The Basic Problem	25
	The Three Concepts of Probability Included in This Study	27
	(1) The Points of A Finite Sample Space	28
	(2) The Probability of A Simple Event for A Finite Sample Space	29
	(3) Quantification of Probabilities	32
	Summary of the Problem	33
	The Hypotheses and Questions	34
III.	PROCEDURES AND DESIGN OF THE STUDY	36
	Population	36
	Subjects	43
	The Pilot Studies	44
	The Tests	46
	Administering the Tests	50
	Design of the Study	54
IV.	RESULTS OF THE RELIABILITY STUDIES	60
	Description of the Statistics Used in the Reliability Studies	60
	Reliability Studies of Test I (Sample Space)	63
	Reliability Studies of Test II (Probability of a Simple Event)	71
	Reliability Studies of Test III (Quantification of Probabilities)	78

CHAPTER		PAGE
IV.	Continued	
	Reliability Studies of Subtest I-A (Sample Space: Simple Counting)	85
	Reliability Studies of Subtest I-B (Sample Space: Combinations)	90
	Reliability Studies of Subtest II-A (Probability: Simple Counting)	95
	Reliability Studies of Subtest II-B (Probability: Combinations)	100
	Summary	105
V.	ANALYSIS OF THE DATA	107
	Part 1: The Testing of Hypotheses	107
	Part 2: Correlation Studies	129
	Part 3: Analysis of Incorrect Responses on the Test Items	138
	Subtest I-A	138
	Summary of Errors on Subtest I-A . . .	146
	Subtest I-B	147
	Summary of Errors on Subtest I-B . . .	153
	Subtest II-A	153
	Summary of Errors on Subtest II-A . .	160
	Subtest II-B	161
	Summary of Errors on Subtest II-B . .	168
	Test III	168
	Summary of Errors on Test III	180
VI.	SUMMARY AND CONCLUSIONS	181
	Results of the Study	183
	Part 1: Results of the Tests of Hypotheses . .	184
	Part 2: Results of the Correlation Studies . .	186
	Part 3: Results of the Analyses of Incorrect Responses on the Test Items	189
	Conclusions and Implications	191
	Status of the Concepts	192
	Implications	192
	The Relationship of the Overall Performance on the Tests With the Factors of I.Q., Sex and Grade	195
	Suggestions for Further Study	196

APPENDIX PAGE

A Probability Tests I, II and III, Directions for Administering the Tests, and Answer Key 198

B Inter-item Correlations for Subtests I-A and II-A for Grades Four, Five, Six and Seven 224

C Inter-item Correlations for Subtests I-B and II-B for Grades Four, Five, Six and Seven 229

D Scores of 528 Children on the California Test of Mental Maturity, Stanford Arithmetic Achievement Test, and the Probability Tests 234

BIBLIOGRAPHY 259

LIST OF TABLES

TABLE		PAGE
1	Population by Grade, Dates of Administration and Form of the California Test of Mental Maturity Used for Selection of the Population	38
2	Frequencies of Total I.Q.'s for Each Grade and Total Population	39
3	Frequencies of Grouped Total I.Q.'s for the Population at Each Grade Level and the Total Population	41
4	Mean Scores and Standard Deviations of Total I.Q.'s for the Population at Each Grade Level and Total Population	41
5	Number of Children in Each of the Twenty-four Subclasses of the Total Population	43
6	Schematic Representation of the Design of the Study .	55
7	MANOVA Table	58
8	Frequency Distributions of Total Scores on Test I for Grades Four, Five, Six and Seven	66
9	Mean Scores and Standard Deviations for Test I	66
10	ANOVA and Reliability Table for Test I, Grade 4 . . .	67
11	ANOVA and Reliability Table for Test I, Grade 5 . . .	67
12	ANOVA and Reliability Table for Test I, Grade 6 . . .	68
13	ANOVA and Reliability Table for Test I, Grade 7 . . .	68
14	Item Analysis for Test I for Grades Four, Five, Six and Seven.	69
15	Frequency Distributions of Total Scores on Test II for Grades Four, Five, Six and Seven	73
16	Mean Scores and Standard Deviations for Test II . . .	73
17	ANOVA and Reliability Table for Test II, Grade 4 . .	74

TABLE		PAGE
18	ANOVA and Reliability Table for Test II, Grade 5 . . .	74
19	ANOVA and Reliability Table for Test II, Grade 6 . . .	75
20	ANOVA and Reliability Table for Test II, Grade 7 . . .	75
21	Item Analysis for Test II for Grades Four, Five, Six and Seven	76
22	Frequency Distributions of Total Scores on Test III for Grades Four, Five, Six and Seven	80
23	Mean Scores and Standard Deviations for Test III . . .	80
24	ANOVA and Reliability Table for Test III, Grade 4 . .	81
25	ANOVA and Reliability Table for Test III, Grade 5 . .	81
26	ANOVA and Reliability Table for Test III, Grade 6 . .	82
27	ANOVA and Reliability Table for Test III, Grade 7 . .	82
28	Item Analysis for Test III for Grades Four, Five, Six and Seven	83
29	Frequency Distributions of Total Scores on Subtest I-A for Grades Four, Five, Six and Seven	86
30	Mean Scores and Standard Deviations for Subtest I-A .	86
31	ANOVA and Reliability Table for Subtest I-A, Grade 4 .	87
32	ANOVA and Reliability Table for Subtest I-A, Grade 5 .	87
33	ANOVA and Reliability Table for Subtest I-A, Grade 6 .	88
34	ANOVA and Reliability Table for Subtest I-A, Grade 7 .	88
35	Item Analysis for Subtest I-A for Grades Four, Five, Six and Seven	89
36	Frequency Distributions of Total Scores on Subtest I-B for Grades Four, Five, Six and Seven	91
37	Mean Scores and Standard Deviations for Subtest I-B .	91
38	ANOVA and Reliability Table for Subtest I-B, Grade 4 .	92

TABLE		PAGE
39	ANOVA and Reliability Table for Subtest I-B, Grade 5 .	92
40	ANOVA and Reliability Table for Subtest I-B, Grade 6 .	93
41	ANOVA and Reliability Table for Subtest I-B, Grade 7 .	93
42	Item Analysis for Subtest I-B for Grades Four, Five, Six and Seven	94
43	Frequency Distributions of Total Scores on Subtest II-A for Grades Four, Five, Six and Seven	96
44	Mean Scores and Standard Deviations for Subtest II-A .	96
45	ANOVA and Reliability Table for Subtest II-A, Grade 4.	97
46	ANOVA and Reliability Table for Subtest II-A, Grade 5.	97
47	ANOVA and Reliability Table for Subtest II-A, Grade 6.	98
48	ANOVA and Reliability Table for Subtest II-A, Grade 7.	98
49	Item Analysis for Subtest II-A for Grades Four, Five, Six and Seven	99
50	Frequency Distributions of Total Scores on Subtest II-B for Grades Four, Five, Six and Seven	101
51	Mean Scores and Standard Deviations for Subtest II-B .	101
52	ANOVA and Reliability Table for Subtest II-B, Grade 4	102
53	ANOVA and Reliability Table for Subtest II-B, Grade 5	102
54	ANOVA and Reliability Table for Subtest II-B, Grade 6	103
55	ANOVA and Reliability Table for Subtest II-B, Grade 7	103
56	Item Analysis for Subtest II-B for Grades Four, Five, Six and Seven	104
57	Summary Table of Reliability Coefficients of the Probability Concept Tests and Subtests for Grades Four, Five, Six and Seven	106

TABLE PAGE

58 Statistics for Regression Analysis with Three Covariates 110

59 Mean Scores of the I.Q. Groups 112

60 MANOVA Table for Hypothesis 1 113

61 Mean Scores of Boys and Girls 116

62 MANOVA Table for Hypothesis 2 117

63 Mean Scores of the Four Grades 120

64 MANOVA Table for Hypothesis 3 121

65 Mean Scores of Interaction, I.Q. x Sex 124

66 MANOVA Table for Hypothesis 4 124

67 Mean Scores of Interaction, I.Q. x Grade 125

68 MANOVA Table for Hypothesis 5 125

69 Mean Scores of Interaction, Sex x Grade 127

70 MANOVA Table for Hypothesis 6 127

71 Mean Scores of Interaction, I.Q. x Sex x Grade 128

72 MANOVA Table for Hypothesis 7 129

73 Correlations Between the Three Scores on the California Test of Mental Maturity and the Scores on the Three Probability Tests for Children in the Four Grades . . 131

74 Correlations Between Scores on Test I and Test II . . 133

75 Correlations Between Scores on Test I and Test III . . 133

76 Correlations Between Scores on Test II and Test III . 133

77 Mean Scores for Subtest I-A and Subtest II-A 135

78 Correlations Between Total Scores on Subtest I-A and Subtest II-A 135

79 Correlations Between Corresponding Items on Subtests I-A and II-A 135

TABLE		PAGE
80	Mean Scores for Subtest I-B and Subtest II-B	137
81	Correlations Between Total Scores on Subtest I-B and Subtest II-B	137
82	Correlations Between Corresponding Items on Subtests I-B and II-B	137
83	Per Cent of Incorrect Responses on Item 1	139
84	Per Cent of Incorrect Responses on Item 2	140
85	Per Cent of Incorrect Responses on Item 3	141
86	Per Cent of Incorrect Responses on Item 4	143
87	Per Cent of Incorrect Responses on Item 5	144
88	Per Cent of Incorrect Responses on Item 6	145
89	Per Cent of Incorrect Responses on Item 7	147
90	Per Cent of Incorrect Responses on Item 8	148
91	Per Cent of Incorrect Responses on Item 9	149
92	Per Cent of Incorrect Responses on Item 10	150
93	Per Cent of Incorrect Responses on Item 11	151
94	Per Cent of Incorrect Responses on Item 12	152
95	Per Cent of Incorrect Responses on Item 13	154
96	Per Cent of Incorrect Responses on Item 14	155
97	Per Cent of Incorrect Responses on Item 15	156
98	Per Cent of Incorrect Responses on Item 16	157
99	Per Cent of Incorrect Responses on Item 17	158
100	Per Cent of Incorrect Responses on Item 18	159
101	Per Cent of Incorrect Responses on Item 19	161
102	Per Cent of Incorrect Responses on Item 20	162
103	Per Cent of Incorrect Responses on Item 21	163

TABLE PAGE

104 Per Cent of Incorrect Responses on Item 22 165

105 Per Cent of Incorrect Responses on Item 23 166

106 Per Cent of Incorrect Responses on Item 24 167

107 Per Cent of Children Who Chose Each Option on Item 25. 169

108 Per Cent of Children Who Chose Each Option on Item 26. 170

109 Per Cent of Children Who Chose Each Option on Item 27. 171

110 Per Cent of Children Who Chose Each Option on Item 28. 172

111 Per Cent of Children Who Chose Each Option on Item 29. 174

112 Per Cent of Children Who Chose Each Option on Item 30. 175

113 Per Cent of Children Who Chose Each Option on Item 31. 176

114 Per Cent of Children Who Chose Each Option on Item 32. 177

115 Per Cent of Children Who Chose Each Option on Item 33. 178

116 Per Cent of Children Who Chose Each Option on Item 34. 179

14

LIST OF FIGURES

FIGURE		PAGE
1	Distributions of Discriminant Scores for I.Q. Groups	114
2	Distributions of Discriminant Scores for Boys and Girls	118
3	Distributions of Discriminant Scores for Grades	122

Abstract

This study examined the status of three concepts, basic to fundamental notions of probability, possessed by children in grades four through seven who had not had any formal learning experiences with topics in probability. The three concepts included in this investigation were: points of a finite sample space; probability of a simple event in a finite sample space; and quantification of probabilities.

The study was carried out during the first semester of the 1967-68 academic year in the Wausau, Wisconsin, Public School System. The population consisted of all children enrolled in grades four through seven for whom a Total I.Q. on the California Test of Mental Maturity was available from the school records. The population included approximately 87% of the total number of children enrolled in grades four through seven in the district in October, 1967. The sample for the study consisted of 528 children randomly selected from the population. The children in sample were stratified into twenty-four subgroups on the basis of sex, three I.Q. ranges and four grades.

Three tests, one for each of the three concepts listed above, were constructed by the writer for use in this study. Each test consisted of a set of items for which the child's responses would indicate if he could apply the concept in a variety of simple experiment and game situations. Each test included a diagram representing the objects used in the experiment or game.

Test I consisted of twelve items on the concept of sample space. The first six items involved only simple counting. The last six items involved simple ideas of combinations.

Test II consisted of twelve items on the concept of probability of a simple event. Each item on Test II presented a lot-drawing situation very similar to the situation presented in the corresponding item on Test I. The first six items on Test II tested the notion of probability of a simple event in which the underlying ideas of sample space involved only simple counting. The last six items tested the notion of probability of a simple event in which the underlying ideas of sample space involved combinations.

Test III consisted of ten items on the concept of quantification of probabilities. Each item presented a game situation in which the child had to decide which of two conditions represented the better probability of success for a specified simple event in one trial. Five of the items presented situations in which the specified event had the same probability of success under both conditions.

The tests were administered as written tests to groups of subjects. The same tests were administered to all subjects, grades four through seven. The items on all tests were scored either right or wrong.

A multivariate analysis of covariance was performed on the results of the three tests. Grade equivalent scores on the three parts of the Stanford

Arithmetic Achievement Test were used as covariates. In a multivariate sense the overall mean performances, adjusted for the covariates, were significantly different ($p<.01$) among I.Q. groups, sex groups and grades. There were no significant interactions.

A univariate analysis of variance was also performed on each dependent variable to determine the level of internal differences for significant overall effects. These results showed:

1. The significant variation among mean vectors for I.Q. groups could be attributed to significant differences among means on all of the probability tests. The mean performances on all three tests ranged from high for the high I.Q. group to low for the low I.Q. group.

2. The variation among the mean vectors for boys and girls could be attributed to the marginally significant mean differences on Test I (sample space) and Test III (quantification of probabilities).

3. The significant variation among mean vectors for grades was due mainly to the significant mean differences on Test I. The mean performances of the children in the four grades ranged from high for the seventh grade to low for the fourth grade over all tests.

Test I was easiest for all grades and Test II (probability of a simple event) was most difficult for all grades. The items on Test II which involved combinations were extremely difficult for all grades.

An analysis of the errors that children made on each of the test items was also performed in an attempt to determine what misconceptions children may have about the concepts tested.

The most significant outcome of this study is that the children demonstrated that they had acquired considerable knowledge about the three concepts of probability under investigation and could apply these concepts on a variety of situations. These children had not received formal training on the notion of probability so their understanding and ability to apply these concepts must have developed as a result of their background, experience and intuition.

The most important implication for educational practice that arises from this study is that since young children acquire some knowledge of probability outside of school, it seems reasonable to assume that some topics of probability would not be too difficult to include in the elementary school program.

Chapter I

BACKGROUND OF THE PROBLEM

Introduction

Probability theory, which had its beginning approximately three hundred years ago, is now considered one of the fastest growing branches of mathematics. Along with the rapidly increasing store of theoretical knowledge has come an ever increasing number of practical applications of this theory. Therefore, probability has become a major topic of interest not only for mathematicians but also for many educators, businessmen, research workers and other members of the general public whose work and lives have been affected by the applications of this mathematical theory.

The eighteenth century French mathematician, Simon Pierre de Laplace, wrote in his treatise on probability,

> We see ... that the theory of probability is at bottom only common sense reduced to calculation; it makes us appreciate with exactitude what reasonable minds feel by a sort of instinct without being able to account for it ... it is remarkable that this science of probability, which originated in the consideration of games of chance, should have become the most important object of human knowledge. [1]

---

[1] E. T. Bell, Men of Mathematics (New York: Simon and Schuster Co., 1937), p. 73

Undoubtedly there are some who would challenge the statement that probability is, or in fact, ever was the most important object of human knowledge. However, there can be no doubt that for the past one hundred fifty years, the study and applications of probability theory have maintained a lofty and highly significant place in the development of human knowledge. The very important role that probability plays in our highly technical world of today is apparent if one considers the ubiquitous applications of this theory. For example, probability is used in the determination of insurance rates; research in the physical, medical and social sciences; the development of military strategy; and as a key to the solution of many complex problems in most major industries in our country. In fact, it is very difficult to think of any profession in which probability theory is not applicable.

In recent years, probability and statistical theory has also gained a greater academic prominence. Many colleges and universities have introduced new courses in probability and statistics, and some of the larger schools have created new departments of statistics, offering undergraduate and graduate majors in the field of statistics and probability. In addition to the increased emphasis on the study of probability theory at the college and university level, recommendations have also been made to include the study of probability and statistics at the high school, the junior high school and even the elementary school level.

The following paragraphs review some of the recommendations, and the rationale for making such recommendations, for including probability in the K-12 mathematics curriculum. Also included is a brief survey of some of the materials available for teaching this topic at various grade levels.

Summary of Recommendations for Teaching Probability in the Secondary and Elementary Schools

The Commission on Mathematics, in its 1959 report, recommended that a course in probability and statistics be included as part of the high school mathematics curriculum. The report states,

> So great is the current scientific and industrial importance of probability and statistical inference that the Commission does not believe valid objections based on theorectical considerations can be offered to its inclusion in the curriculum. [2]

The Commission did feel a valid objection to this proposal might be a lack of suitable material to use for such a course. For this reason the Commission prepared a textbook that was considered appropriate for use in a one-semester high school course. [3]

Since the publication of the 1959 Commission Report several other textbooks on probability and statistical inference, considered suitable

---

[2]Commission on Mathematics College Entrance Examination Board, Program for College Preparatory Mathematics (New York: College Entrance Examination Board, 1959), p. 32

[3]Commission on Mathematics College Entrance Examination Board, Introductory Probability and Statistical Inference, An Experimental Course (Princeton: Educational Testing Service, 1959)

for use in the high school, have been published. [4] It is important to note that many secondary schools now include a course in probability as part of their regular mathematics course offerings.

Some attention has also been given to the topic of probability in several of the newer mathematics programs for junior high school and in a few of the more recent recommendations for improving the junior high mathematics curriculum. It is important to note that these recommendations and the available text materials assume that the pupils already have an intuitive notion of probability and chance events.

Willoughby recommends that the topic of probability be included in the junior high school mathematics curriculum. He says,

> The subject of probability has considerable significance in the world today. It is of interest to people of all ages, including junior high school pupils. One of the best ways to arouse interest in mathematics is to present some topic which arrests the attention of the pupils. Probability is such a topic. [5]

He goes on to say that junior high pupils will have various ideas about probability from their past experiences and that it is desirable to find out what pupils already know before beginning such a unit.

---

[4] Howard Fehr, L. Bunt, G. Grossman, An Introduction to Sets, Probability and Hypothesis Testing (Boston: Heath Co., 1964); Samuel Goldberg, An Introduction to Probability (Englewood Cliffs, New Jersey: Prentice Hall Co., 1960); Frederick Mosteller, R. Rourke, G. Thomas, Probability: A First Course (Reading, Massachusetts: Addison Wesley Co., 1961)

[5] Stephen S. Willoughby, Contemporary Teaching of Secondary School Mathematics (New York: John Wiley and Son, 1967), p. 169

Van Engen, et al., include a unit on probability in their 8th grade textbook.[6] In the teacher notes for the first lesson the authors say,

> The approach in this lesson (Events and their Probabilities) is somewhat intuitive and informal because we assume that the students have had some experience in determining the likelihood that some chance event will or will not occur.[7]

The report of the Cambridge Conference on School Mathematics recommends that probability be included in the K-12 mathematics curriculum. The report includes a special section in the appendices which lists the reasons for including probability in the curriculum and also includes a brief outline of the topics to be included in the K-12 program. In its proposals for the 7-12 mathematics curriculum the Cambridge Committee says this about probability,

> Both programs include several rounds of probability. It is presupposed that the student will have had as a first round an intuitive concept of the probability of an event from the pre-mathematical material in the lower grades. Based upon this, the second round studies finite events using the techniques of algebra. The third round is tied in with the study of infinite sequences and deals with countable event sets, while the fourth round treats the continuous case using calculus.[8]

In addition to the material mentioned above, units on probability are found in several junior high textbooks such as the material

---

[6]Henry Van Engen, M. Hartung, H. Trimble, E. Berger, R. Cleveland, Seeing Through Mathematics, Part 2, Book 2 (Chicago: Scott, Foresman and Company, 1961), pp. 449-484

[7]Henry Van Engen, et al., Teaching Guide, Seeing Through Mathematics, Part 2, Book 2 (Chicago: Scott, Foresman and Company, 1963), p. 260

[8]Cambridge Conference on School Mathematics, Goals for School Mathematics (Boston: Houghton Mifflin Co., 1963), pp. 48-49

published by SMSG, the Maryland Project and other commercial companies.

Recommendations for including probability in the elementary school mathematics program and suggested activities for teaching this topic are beginning to appear in the literature.

Page [9] suggests the inclusion of some work with probability in grades K-12 in order to add variety and interest to the mathematics curriculum and also to provide a novel and pleasant source of drill work with some of the more fundamental parts of the curriculum.

In making specific suggestions for probabilistic activities for the elementary school Page says,

> Throughout the elementary grades there are many opportunities to give children a better understanding of standard topics while giving them some ideas about probability. [10]

Page goes on to list several activities that he feels can be used at each of the various grade levels, primary, intermediate, junior high and high school.

Smith [11] makes several suggestions for activities in probability that can be included in grades 4 - 6. He claims that the types of activities he suggests, dealing with the probability of a simple event, can be readily used in the elementary classroom.

---

[9] David A. Page, "Probability," The Growth of Mathematical Ideas, Grades K-12, Twenty-fourth Yearbook of the National Council of Teachers of Mathematics (Washington, D. C.: The Council, 1959), pp. 229 - 271

[10] Ibid., p. 232

[11] Rolland R. Smith, "Probability in the Elementary School," Enrichment Mathematics for the Grades, Twenty-seventh Yearbook of the National Council of Teachers of Mathematics (Washington, D.C.: The Council, 1963) pp. 127 - 133

Deans [12] includes the principles of probability among the basic mathematical concepts that should be stressed in the elementary school.

Osborn, et al., recommend that statistics and probability be included in the elementary program. The authors make the following statements about the type of activities that should be included.

> Many statistical processes, techniques, and understandings are well within the grasp of the elementary school child. Since we more and more frequently try to communicate with each other by associating numbers with physical phenomena, the child at an early age should be introduced to those phases of statistics which are within his grasp. It seems relatively certain that our approach to statistics in the elementary grades should center around experiences in the child's ever expanding environment, and that his activities should include collection, organization, and interpretation of data and the more sophisticated arts of inference and decision-making. [13]

Osborn also states that the subject of probability is very closely associated with decision-making which suggests that the elementary child must have an understanding of probability before he can engage in recommended activities involving inference or decision-making.

In an unpublished report of a summer writing team, [14] the Cambridge Committee includes a chapter listing specific activities that can be used to teach some of the basic concepts of probability to

---

[12] Edwina Deans, Elementary School Mathematics: New Directions, (Washington, D. C.: U. S. Office of Education, 1963) p. 114

[13] Roger Osborn, M. W. DeVault, C. C. Boyd, W. R. Houston, Extending Mathematical Understanding (Columbus, Ohio: Merrill Company, 1961), p. 194

[14] Cambridge Conference on School Mathematics, Reports of 1965 Summer Study, Unpublished mimeograph report (May, 1966), Section II - Chapter 4

children in the elementary school. The following statement from this report summarizes the Committee's reasons for suggesting the study of probability in the elementary school.

> Today, probability is one of the most widely used branches of mathematics, not only in various vocations, but in the everyday life of the 'Man in the Street.' The ordinary citizen is constantly bombarded with statistics about toothpaste, automobile accidents, the probability that there is a connection between smoking and various kinds of illnesses, the probability that candidate A is going to win an election, etc.
>
> As well as being useful in the real world, probability is an interesting and exciting means of getting children to practice arithmetic. It is also a good mathematical model of the real world, and offers children considerable practice in creating mathematical models with approximate reality.
>
> All of these reasons seem to point to the early teaching of some probabilistic concepts in the elementary grades. Certainly, a considerable amount of probability should be learned by all pupils before some discontinue their formal mathematical education. A further reason for the early introduction of probability into the curriculum is that many people have the feeling that mathematics studies only exact data and exact numbers - - probability will give the feeling of studying distributions and uncertainties before the pupils become overly enamoured with 'getting the exact answer.' 15

Each of these recommendations for including the study of probability in the elementary school suggest some specific ideas that can be used as the basis for class activities but do not specify when these topics should be introduced, except in a very general manner. The Cambridge Committee 1965 Summer Study Report does list specific activities that might be appropriate for each of the grades 3, 4, 5 and 6, but also states that experience may show that it might be

---

[15] Ibid., p. 1

better to delay some of the topics for a grade or two.

It is apparent from the types of activities proposed for the elementary school that the authors of these recommendations assume that the children already have some intuitive notions about certain basic concepts of probability such as identifying all of the outcomes in a sample space, assigning probabilities to simple events, and recognizing when events are equally likely.

Short units on probability are found in three of the newer elementary arithmetic textbook series. The series Sets and Numbers [16] has a 6-page unit in Book 4, a 5-page unit in Book 5, and a 7-page unit in Book 6. The series Modern School Mathematics [17] has a 4-page unit in Book 5 and a 6-page unit in Book 6. The revised edition of the series Seeing Through Arithmetic [18] has a 10-page unit in Book 6. An examination of these units reveals that these authors also presume that the children already have some knowledge about the basic probability concepts mentioned above.

---

[16] Ernest R. Duncan, Lelon R. Capps, Mary P. Doliciani, W. G. Quast, Marilyn Zweng, Modern School Mathematics Structure and Use, Books 5 and 6 (Boston: Houghton Mifflin Company, 1967), Book 5, pp. 252-255; Book 6, pp. 226-229, 250-251

[17] Patrick Suppes, Dolly Kyser, Catherine Braithwaite, Sets and Numbers, Books 4, 5 and 6 (New York: Singer Company, 1966), Book 4, pp. 316-321; Book 5, pp. 294-298; Book 6, pp. 314-321

[18] M. Hartung, H. Van Engen, E. G. Gibb, J. Stochl, L. Knowles, R. Walch, Seeing Through Arithmetic 6 (Chicago, Scott, Foresman and Company, 1968), pp. 289-298

The most extensive set of text material for teaching probability in the elementary school that is currently available is a set of two booklets, Probability for Primary Grades and Probability for Intermediate Grades[19] written by the School Mathematics Study Group. These materials do not presume any prior knowledge of probability concepts and each of the aforementioned basic concepts is developed through a series of suggested class activities and written exercises. These materials are not graded but it is suggested by the authors that some of the units in the first book are appropriate for kindergarten and first grade with the latter units being more appropriate for second and third grades.

Need for Research

As was pointed out in the preceding section, some mathematics educators are recommending that the topic of probability should be included in the elementary school arithmetic program. The majority of the activities suggested by these educators, and the units on probability included in some of the newer elementary arithmetic textbooks, are based on the assumption that children already possess some basic concepts of chance and are able to apply these concepts in simple probabilistic situations. However, none of the recommendations or text materials list what must be considered as the minimum requirements that

---

[19] School Mathematics Study Group, Probability for Primary Grades and Probability for Intermediate Grades (Palo Alto, California: Stanford University, 1965)

the children must possess before such activities can be used successfully in the elementary grades.

Because of the increasing importance of probability in the real world it is reasonable to assume that in the near future this topic will receive considerable attention from curriculum workers and elementary arithmetic textbook writers. It is important that research evidence be available to help these workers answer pertinent questions regarding the placement of this topic in the elementary school curriculum. Grade placement of probability topics varies greatly in materials that are currently available. For example, activities involving sample space and probability of a simple event are introduced as early as the primary grades in the SMSG booklets on probability while similar ideas are not introduced until the fifth or sixth grade in other textbooks such as Modern School Mathematics Structure and Use, Books 5 and 6, Houghton Mifflin Company and Seeing Through Arithmetic 6, Scott, Foresman and Company. Therefore, it seems important that some preliminary investigation be conducted to assess the status of basic probability concepts in elementary school children.

Very little research has been done on the placement of new topics, such as probability, in the elementary school mathematics curriculum. However, it is generally agreed that experimental evidence which may suggest the optimal time for the introduction of new concepts would undoubtedly help to improve the curriculum. Also, little research has been conducted to determine when the basic concepts of probability, which are assumed to exist in the minds of elementary school children, seem to emerge. The studies of Piaget [20] and Leake [21] indicate that children do acquire some concepts of probability outside of school, but more evidence is needed. It is desirable to ascertain whether children have a good understanding of the basic concepts of probability; such as, the points of a sample space and the probability of a simple event, which includes the ability to apply these concepts in a wide variety of situations, before the children can be expected to pursue a more systematic treatment of the ideas of probability.

This study was devised to provide some evidence concerning the development of probability concepts with children in grades four through seven. It is an attempt to determine when three basic concepts of probability begin to emerge in young children as a result of their background, experiences and intuition. The subjects are children who have not had any formal learning experiences with topics in probability. The subjects were categorized into groups by I.Q., sex and grade. Grade placement scores on a standardized arithmetic achievement test were also obtained. The study includes an analysis of the relationships between these factors and the level of proficiency the subjects display in applying the three basic probability concepts in game situations.

The significance of this study for education is evident. If it

---

[20] Jean Piaget, Barbel Inhelder, LaGenése de l'Idée de Hasard chez l'Enfant (Paris: Presses Universitaires de France, 1951)

[21] Lowell Leake, The Status of Three Concepts of Probability in Children of 7th, 8th and 9th Grades, Unpublished Ph.D. Dissertation (University of Wisconsin, 1962)

is found that young children do acquire ideas of probability outside of school and can intuitively apply these ideas to problem situations, this would suggest that this topic is not too difficult to teach to elementary school children. It is important to try to find what intuitive notions children of various ages can be expected to already know and what concepts of probability might be most appropriate to teach at each grade level.

It is also important to know the relationship between I.Q., arithmetic achievement scores and success on the probability tests used in this study. The relationship between these factors and the success in applying basic ideas of chance may indicate to curriculum builders the appropriateness of including such topics for all children or only for special groups of children in the various grades.

<u>Interpretations of the Term "Probability"</u>

Despite the importance that the study of probability has achieved and the number of eminent mathematicians and philosophers involved in its development, there is no universal agreement on the meaning of "probability."

Historically there are three main interpretations of the word probability.[22] The classic view, attributed to Laplace and DeMorgan, is called the *a priori* definition of probability. This point of view implies that each simple event in a sample space can be assigned a

---

[22] J. Newman, *The World of Mathematics* (New York: Simon and Schuster Company, 1956), pp. 1395 - 1414

number p, $0 \leq p \leq 1$, which represents the degree of certainty that event will occur if one element of the sample space is selected at random. The number p is called the probability of the simple event. The sum of the numbers p assigned to all possible simple events in the sample space must be equal to 1. A certain event is always assigned the number 1 and an impossible event is always assigned the number zero.

A second interpretation of probability is that of an intuitive relation which has degrees, but is not capable of being analyzed numerically. That is, although the degrees of this "probability relation" exist, they are not measurable. A statement such as, "It is probable that I would have had a good time if I had gone to the party," describes a situation in which the use of the word probable fits under this second interpretation.

The third definition of probability, sometimes referred to as the <u>a posteriori</u> or frequency definition of probability, is the interpretation that is most widely accepted. The determination of a probability by this method is empirical or experimental. An experiment or trial is repeated n times, noting the number of successes s. The number $\frac{s}{n}$ is called the probability of success for future trials of the same nature. Technically this number is incorrect since it is impossible to consider all possible trials. Thus the number $\frac{s}{n}$ is an approximation to the probability of success which is defined as the $\lim_{n \to \infty} \frac{s}{n}$.

Interpretation of Probability to be Used in this Study

The theory of probability can be considered as the study of mathematical models of chance events. To develop any mathematical theory it is necessary to first define the class of things about which this theory will deal. The class of interest for a particular theory is often called a universe or space. In probability theory the class of things with which one deals is a set of experiments involving chance events and the outcomes of such experiments. Therefore, the first task in developing a precise and useful theory of probability is specifying all possible outcomes of an experiment. The set of all possible outcomes is called a sample space of the experiment and the elements of this set are called points of the sample space.

The main interest in specifying the points of a sample space is not in the points themselves, but rather in the probabilities with which these outcomes can occur. Therefore, the second task in a systematic development of a theory of probability is assigning a number, called the probability of the outcome, to each point in the sample space. Under the *a priori* interpretation of probability these probabilities can be assigned to the points arbitrarily, but must satisfy the following two conditions:

1. The probability assigned to each point is a non-negative number p with $0 \leq p \leq 1$.
2. The sum of the probabilities assigned to all of the points in the sample space is 1.

The discussion of probability in this study will be limited to the interpretation described above in which the two underlying concepts are points of a sample space and probability of a simple event. An understanding of these concepts is essential before a systematic study of the more complex notions of probability can be attempted.

Related Research

The most prominent of the studies dealing with the development of probability concepts in young children is the series of experiments conducted by Jean Piaget and Barbel Inhelder.[23] In this study the authors presented a series of tasks to children between the ages of 7 - 14 in order to determine the child's conception of chance phenomena. The study consisted of three parts. The first part dealt with the subject's notion of randomization, the idea of uniform and normal distributions resulting from random movements of physical objects and the relation between chance events and nonrandom events. The second part studied the subject's conception of probability notions by having him predict events in lot-drawing, coin tossing and drawing colored counters out of a bag. Another aspect of this part included the quantification of probabilities in situations similar to the above. The third part was designed to study the subject's understanding of combinations, permutations and ordered arrangement of objects. In each part of the study the situations were presented to the subjects using demonstrations with concrete objects. The subjects were interviewed by the authors, and the subjects' actions and responses were

[23]Piaget, et al., op. cit.

recorded as they performed the tasks and answered questions during the experiments.

In his conclusions, Piaget classifies three stages in the development of the ideas of chance. Although Piaget considers the order in which the stages emerge to be invariant, he points out that the age at which a given stage appears may vary considerably. Flavell summarizes Piaget's comments regarding the relationship between chronological age and developmental stages as follows:

> ... Piaget readily admits that all manner of variables may affect the chronological age at which a given stage of functioning is dominant in a given child: intelligence, previous experience, the culture in which the child lives, etc. For this reason, he cautions against an over-literal identification of stage with age and asserts that his own findings give rough estimates at best of the mean ages at which various stages are achieved in the cultural milieu from which his subjects were drawn. [24]

The first stage, between 7 and 8 years of age, is what Piaget calls the pre-operational stage. This stage is characterized by an indifference between chance and certainty because the child does not as yet have the intellectual operations necessary to recognize those events which are certain, much less those which are uncertain. The pre-operational child is not able to predict events based on probabilities or quantify probabilities because he has not yet acquired an understanding of proportions which Piaget claims is fundamental to

---

[24] J. H. Flavell, The Developmental Psychology of Jean Piaget (Princeton, New Jersey: D. Van Nostrand Company, Inc., 1963) p. 20

the understanding of probability.

The second stage, ages 8 - 11, Piaget defines as the concrete operation's stage. During this stage the child begins to understand the difference between chance and certainty but is not able to distinguish the degrees of chance based on probability notions. The child does understand the notion of random mixtures but as yet has not acquired the idea of the "law of large numbers" which is necessary to predict the shape of a random distribution. He is able to quantify probabilities in a limited way but often makes the error of basing his judgements on the absolute number of objects that are used rather than comparing the number of objects which represent successes to the total number of objects used in the experiment.

The third stage, ages 11 and over, is described as the formal-operational stage. At this stage the child is able to work very well with chance events. He is able to generalize the ideas of uniform random distributions, he is able to quantify probabilities and has the capacity to compute combinations, permutations and arrangements in a systematic manner.

This study of Piaget's, like many of his other works, is a valuable contribution to the study of human development. However, this work like many of the others lacks a number of things that are often considered fundamental in reporting the results of a research study. The testing and interview procedures are not precisely described. The report does not include sample size, method of selection, or background of the subjects. There is no coefficient of reliability reported for the evaluation instrument used in the study. It is not clear whether the same tasks or evaluation procedures were used on all subjects. The organization and analysis of data is very limited and in most cases the results cited are merely subjective judgements with no attempt to provide quantitative information on the data.

Only a few studies have been conducted in an attempt to test Piaget's findings on the development of the concept of probability. Other experiments have been conducted which in some way deal with the development of this concept, but often this is a secondary issue rather than the main question of the study. A few of the most important of these studies are summarized below.

Yost, Siegel and Andrews [25] tested Piaget's conclusion that children before the age of 7 are not able to utilize the concept of probability. They criticized Piaget's methods for the heavy reliance on verbal skills, confounding color preference with color expectation, lack of randomization of the concrete memory aids and not providing a statistical analysis of the results. The method used by Yost compared a Piaget method with a "decision method" using the most and least preferred colors of the subjects, objects randomly distributed in transparent plastic boxes and rewards as incentives for correct responses. The test was to select the box that would give the better chance of winning in a lot-drawing situation with different proportions of winning and losing counters. This study showed that young children, 4 - 6 years of age, do demonstrate some understanding of probability

[25] Patricia A. Yost, A. Siegel, J. Andrews, "Nonverbal Probability Judgements by Young Children," Child Development, 33 (1962), pp. 769-780

under appropriate conditions. In this study, scores obtained by the decision method were significantly greater than scores obtained by the Piaget method.

Pire [26] used the multiple choice items of a French intelligence test relative to chance or probability and analyzed the results obtained from a large number of subjects in order to compare them with the results of Piaget's study. His results show a trend similar to the developmental growth reported by Piaget, but reported great individual differences at all levels. Pire showed that general intelligence is a source of variation as well as sex. Boys had higher scores than girls at all age levels. Piaget did not include sex or general intelligence as variables in his study.

Davies [27] analyzed the data obtained from 112 subjects between the ages of 3 - 9 years. Two tests were administered, one verbal and one non-verbal, in which probabilities of $\frac{1}{5}$ and $\frac{4}{5}$ were used. Davies reported findings consistent with Piaget's for the pre-operational child. The study showed that non-verbal behavior of event probability appears earlier than verbalization of the concept. The study also showed that there were no significant differences between sexes which is contrary to the findings of Pire.

---

[26] G. Pire, "Notion du Hasard et Development Intellectual," Enfrance, (1958), pp. 131 - 143

[27] C. M. Davies, "Devélopment of the Probability Concept in Children," Child Development, 36 (September, 1965), pp. 779 - 788

Leake[28] studied the status of three concepts of probability in 72 seventh, eighth and ninth grade pupils. This study was not intended to test the results of Piaget but rather to determine the status (without any formal learning experience) of some of the basic concepts of probability and to test the significance of the variables of school, grade, mathematical ability and sex on the acquisition of the concepts. Leake used a repeated measures analysis of variance design to analyze the data. He reported significant F ratios for the main effects of grade, level of achievement and concepts. He also found a significant inter-action between the level of mathematical achievement and the concepts. Leake concluded that mental age is more important than chronological age in the acquisition of the concepts. The three ten-item tests that Leake used in the study relied on a high level of verbal skill. However, there may be no serious objection to this at the junior high level. Only two schools from a large city school district were used in the study so the sample cannot be considered a random sample of the junior high school population in the district. Leake did not claim his sample was a random sample of the district nor did he attempt to generalize his results. Leake did not include "low-ability" students in his sample, assuming that these students would not know very much about the concepts of probability under investigation. Since the subjects used in the study had not had any formal learning experiences in probability the study was to

---

[28]Leake, op. cit.

assess the status of these concepts based on the subjects' past experiences an intuition. Therefore, there was no basis for assuming that low-ability students would perform differently than other students on the tests used to evaluate the status of the concepts included in the study.

Cohen [29] has conducted a number of experiments dealing with "subjective probability," the use of the concept of probability in risk and gambling situations, and the idea of independence. These studies were not specifically intended to study the development of the concept of probability but the reported results do not differ significantly from the findings of Piaget.

Cohen's results show that young children are greatly influenced by subjective preference, ideas of fairness, and superstitious behavior, in dealing with situations involving probability. Cohen states that his results show that the concept of probability, and independence of events, does not develop before the ages of 14 - 16 or older.

Gratch [30] studied the idea of independence with first, third,

---

[29] John Cohen, C. E. M. Hansel, Risk and Gambling (London: Longmans, Green, 1956); John Cohen, "Subjective Probability," Scientific American, 197 (November, 1957), pp. 128-138; J. Cohen, E. J. Dearnaley, C.E.M. Hansel, "Measures of Subjective Probability," British Journal of Psychology, 48 (1957), pp. 271-275; J. Cohen, Chance, Skill and Luck: the Psychology of Guessing and Gambling (Baltimore: Penguin Books, 1960)

[30] G. Gratch, "The Development of the Expectation of the Non-independence of Random Events in Children," Child Development, 30 (1959), pp. 217-227

fifth and sixth grade subjects. He does not attempt to explain the actions of the younger children but says that the reactions of the 11 year old children is consistent with Piaget's view on "morally realistic" thinking of children, where they do not distinguish between moral rules (fairness) and physical laws based on probability ratios.

Kass[31] studied the reactions of 42 subjects, 6, 8 and 10 years old, in a "pay to play" situation. Different probabilities of payoff, $\frac{1}{1}$, $\frac{1}{3}$ and $\frac{1}{8}$, were used, but the expected return was 0 for all experiments. Kass did not find any significant differences for age but did find that boys tended to pick the situations with low probability, $\frac{1}{8}$, more often, while girls picked the higher probabilities, $\frac{1}{3}$ or $\frac{1}{1}$.

Ross and Levy [32] did an extensive study on what they call the "maturity of chances" effect, also called the "gamblers fallacy" or "negative recency" effect, which is the tendency to prefer alternate predictions to the occurring run of events. This study used fifth grade, eighth grade, tenth grade and adult subjects. The results showed that the tenth grade and adult subjects displayed a greater "maturity of chances" effect. That is, these subjects made more predictions as if the random events were dependently related over a short

---

[31] Norman Kass, "Risk and Decision Making as a Function of Age, Sex, and Probability Preference," Child Development, 35 (1964), pp. 577 - 582

[32] B. M. Ross, N. Levy, "Patterned Predictions of Chance Events by Children and Adults," Psychological Reports, 4 (1958), pp. 87-126

series of repeated outcomes. This study contradicts the findings of Piaget and Cohen with respect to this concept for subjects in the fifth and eighth grades (concrete-operational and formal-operational stages).

Stevenson, Weir, Ziger, Messick, Solley and others [33] have conducted many studies dealing with probability learning in young children but these studies are not primarily concerned with the status of the concept of probability in the subjects. These studies concern themselves with the application of subjective judgements in risk situations or learning experiences which include some aspects of probability as it applies to a specific task.

Several other studies [34] report the results of experiments with the teaching of probability in the primary and intermediate grades. However, none of these studies have attempted to determine the status of the concept of probability in the subjects before the units on probability were presented in the classroom.

---

[33] S. J. Messick, C. M. Solley, "Probability Learning in Children: Some Exploritory Studies," Journal of Genetic Psychology, 90 (1957), pp. 23-32; H. W. Stevenson, M. W. Weir, "The Role of Age and Verbalization in Probability Learning," American Journal of Psychology, 76 (1963), pp. 299-305; H. W. Stevenson, E. F. Zigler, "Probability Learning in Children," Journal of Experimental Psychology, 56 (1958), pp. 185-192; M. W. Weir, "The Effects of Age and Instruction on Children's Probability Learning," Child Development, 33 (1962), pp. 729-735.

[34] Ralph Ojemann, E. J. Maxey, B. C. Sinder, "Effects of A Program of Guided Learning Experiences in Developing Probability Concepts at 3rd Grade," Journal of Experimental Education, 33 (1965), pp. 321-330; Ralph Ojemann, et al., "Effects of Guided Learning Experiences in Developing Probability Concepts at the Fifth Grade Level," Perceptional and Motor Skills, (1965), pp. 415-427; J. D. Wilkinson, O. Nelson, "Probability and Statistics: Trial Teaching in 6th Grade," Arithmetic Teacher, 13 (1966), pp. 100-106

Chapter II

THE PROBLEM

The Basic Problem

The problem under investigation is to examine the status of three basic concepts of probability with children in grades four through seven. Although this study is not intended to replicate Piaget's experiments dealing with the concepts of probability, his work motivated the conception and design of this study. Piaget found that the young children he worked with did acquire some ideas about probability outside of school. Therefore it seems reasonable to hypothesize that children, who have not had any formal learning experiences with the topic of probability, can intuitively apply certain basic concepts of probability in problem situations.

This hypothesis is further supported by the findings of Leake in a status study of probability concepts with children in grades seven through nine. Leake concludes,

> Without any doubt, the students had acquired considerable knowledge and ability to deal with problems about probability on the level of the three elementary concepts. This acquaintance was not due to being taught probability concepts, but must have come from their everyday experiences and exposures in growing....[35]

Other studies cited in Chapter I also concluded that young children can intuitively apply certain probabilistic ideas. However,

---

[35]Leake, op. cit., p. 50

a review of the literature indicates that no study, other than the work of Piaget, has been reported that deals specifically with the status of probability concepts with children of ages nine through thirteen. This is the period which Piaget claims is the formative period during which children somehow acquire fundamental knowledge of probability without formal training. Also, the activities that are being suggested for use in the elementary arithmetic program are recommended for use with children in this age group. Because of the emphasis now being directed toward including probability as part of the intermediate grades' arithmetic program, this study is primarily concerned with the performances of children in the grades four, five, six and seven when applying three concepts of probability to game situations. Although grade seven is generally not considered as an intermediate grade, it was included for comparison purposes.

In his study with junior high school pupils Leake found that level of achievement on a standardized arithmetic test (which he equated with mental age in arithmetic) was a more significant factor than chronological age in determining scores on the tests used.[36] Since standardized arithmetic achievement scores are generally highly correlated with I.Q., it is reasonable to hypothesize that there is a relationship between intuitive knowledge of probability and I.Q.

As previously noted Pire and Kass found that there was a significant difference in the performances of boys and girls on the probability test items included in their studies. However, Davis and

---

[36]*Ibid.*, p. 55

Leake found that the performances of boys and girls did not differ significantly. Of course these studies varied considerably with respect to age of subjects, type of test item, setting and so on. Nevertheless they were all concerned with the subject's ability to apply fundamental concepts of probability to problem situations. Since the results of these studies are contradictory the hypothesis that sex is a significant factor in the acquisition of probability concepts needs further study.

For this study it was decided to partition the subjects on the basis of grade in school, sex and three I.Q. groups. The three I.Q. ranges used to partition the population are: 71-104, 105-113 and 114-144. Approximately one-third of the population for the study was included in each of the three groups.

The Three Concepts of Probability Included in this Study

The study of probability involves many fundamental concepts such as: a sample space for an experiment, simple and compound events, probability of an event, mutually exclusive events, probability of the union of two mutually exclusive events, independent events, probability of the intersection of two independent events and so on. Since it was considered impractical to try to include all of these basic concepts of probability in one study it was decided to limit this study to an investigation of only three concepts. This decision was made so that an adequate number of test items could be included for each concept under investigation and still keep the total testing time for each group of subjects within the limits of from 45 to 60 minutes. By limiting the total testing time each group test could be completed in one sitting. Again, this was a practical limitation since the testing was done in seventeen schools and involved more than five hundred subjects.

The three concepts examined in this study are discussed in the following paragraphs. The first two concepts were selected from the many possible candidates because they are considered to be the two basic concepts that must be acquired before a systematic study of probability can be attempted. Also, the majority of authors proposing probabilistic activities to be included in the elementary school mathematics program assume that elementary school children already have some understanding of these two concepts. The third concept is included in this study because it involves an understanding of proportionality which Piaget considers to be a fundamental notion in the study of probability. Flavell, in his summary of Piaget's work, paraphrases this conclusion of Piaget as "Thus, one intellectual achievement indispensible in calculating probabilities appears to be the ability to deal with proportionality." [37]

The three concepts to be investigated are:

(1) The points of a finite sample space.

A sample space for an experiment can be defined as a set, such that each element of the set is an outcome of the experiment and any outcome of the experiment corresponds to exactly one element of the set. The elements of such a set are called sample points. [38] A sample space can be an infinite set or a finite set. For this study the discussion was limited to experiments in which the corresponding sample spaces were

---

[37] Flavell, op. cit., p. 346

[38] Fehr, et. al., op. cit., p. 92

finite sets with fewer than ten elements.

As an example, consider the experiment in which a fair die is rolled, and an outcome of the experiment is defined to be the number on the upper-most face of the die. The set {1,2,3,4,5} is the sample space for this experiment. It is important that whatever constitutes an outcome for a given experiment is clearly defined, for it is possible to have more than one sample space which describes the same experiment. If an outcome of the experiment described above is considered to be whether the number on the upper-most face is odd or even, then the sample space is the set {odd, even} .

A listing or description of all the elements of a sample space may involve more than just simple counting. It may also include an understanding of combinations, permutations or other sophisticated counting techniques.

In this study a subject's understanding of this concept was determined by his performance on a test in which he was asked to list all possible outcomes for a variety of lot-drawing experiments. Since the purpose of this study was to determine if the children had acquired a basic understanding of the concept, the items on the test involved only simple counting and a fundamental notion of combinations.

In the following pages of this report this concept will be referred to as the concept of sample space.

(2) The probability of a simple event in a finite sample space.

An event is defined as a subset of a sample space. A simple event is a subset containing exactly one element of the set.[39] If a sample space for an experiment contains n elements there are 2^n different

---

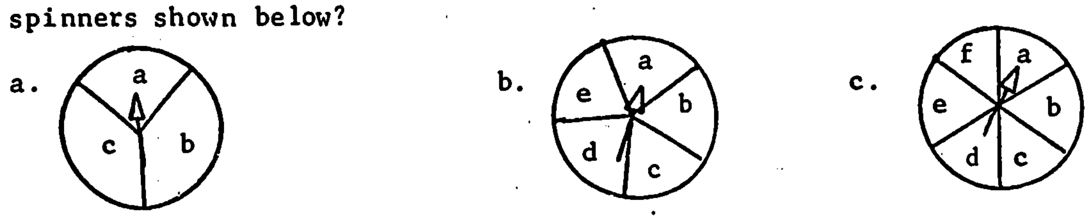
[39]Ibid., p. 96

subsets of the sample space and therefore there are 2^n different events for the experiment. Of course, there are only n simple events. In this study only simple events were considered.

Under the a priori interpretation of probability (which is the interpretation being used in this study), given a sample space for an experiment, each simple event is assigned a non-negative number p called the probability of the event. This assignment is arbitrary but must meet the following conditions of this interpretation: $p_i \leq 1$ and $\sum_{i}^{n} p_i = 1$. If the experiment is set up so that the simple events can be considered equally likely then the intuitive or "natural" choice of probabilities is to assign the same number p to each simple event. That is, if there are n equally likely simple events, each event is assigned the probability $\frac{1}{n}$.

The following examples, selected from elementary textbooks, illustrate the types of exercises involving the probability of an event that are being included in introductory units on probability for elementary school children. The first example is the first exercise on the first page of the unit on probability in Sets and Numbers, Book 4.[40]

What is the probability of getting a when spinning each of the spinners shown below?



The page on which this problem appears represents the children's first formal introduction to probability in this text series. The page

[40]Suppes, et. al., op. cit., p. 316

contains a sample problem in which the terms "probability" and "outcome" are introduced. The children must recognize the sample space for each part of the exercise but they are not asked to list the outcomes. The children must also assume that these figures represent "fair" spinners.

The second example is the third problem on the first page of the unit on probability in Modern School Mathematics, Book 5. [41]

> If there are 7 packages, all wrapped the same way, 5 with candy in them, what is the chance you would pick a package with candy in it?

The page on which this problem appears represents the children's first formal introduction to probability in this text series. One example, similar to the problem above, is explained in some detail at the top of the introductory page to familarize the children with the meaning of the word "chance" as it is being used in the context of the exercises.

To answer this problem correctly (according to the answer given in the answer key) the children must assume that each box has the same chance of being picked. They must also recognize that the events {box with candy} and {box without candy} are not equally likely and have the probabilities $\frac{5}{7}$ and $\frac{2}{7}$ respectively.

Both of the textbook series mentioned above include probability exercises involving combinations in their introductory materials.

In this study a subject's understanding of this concept was determined by his performance on a test in which he was asked, in a variety of ways, to give the probability of winning a simple game in one trial. The items involved simple counting and combinations.

---

[41] Duncan, et. al., op. cit., p. 252

In the following pages of this report this concept will be referred to as the concept of probability of a simple event.

(3) The quantification of probabilities.

The term quantification here refers to a comparison of the probabilities of events; that is, in one trial the occurrence of one event is more probable or less probable than the occurrence of another event or the events are equally probable. It is also possible to compare the occurrence of the same event under different conditions.

One example of an application of this concept is: deciding whether the probability of picking a red ball is greater than, less than, or equal to the probability of picking a blue ball from a box containing 3 red balls and 4 blue balls. It must be assumed that the balls are well-mixed and the draw would be a random selection.

Another example of an application of this concept is: picking the urn which affords the better chance of picking a red chip in one draw, if each of two urns contain some red chips and some blue chips. It must be assumed that the chips in each urn are well-mixed and the draw would be a random selection.

Although this concept is not listed among the basic ideas in mathematical probability textbooks it does involve several notions which are fundamental to an understanding of the ideas of chance. It certainly involves the notion of proportionality. As previously noted Piaget claims that the ability to deal with proportion is essential for an understanding of probability. A conclusion of Piaget, as reported by Flavell, indicates that the preception of young children leads them to make errors in certain probability situations because they can not apply the idea of proportion. Flavell says,

> During middle childhood, the child begins to try to quantify probabilities but repeatedly makes one particular error: he predicts

> solely on the basis of the absolute number of counters with crosses in each collection, rather than in terms of the ratios of these counters to total counters; that is, he seems incapable of reasoning in terms of the proportions in play.[42]

This concept of comparing or quantifying probabilities is included in some of the materials suggested for use in the elementary school. Two examples from Modern School Mathematics, Book 5 serve as examples of how children are expected to demonstrate an understanding of this concept.

> For B = {11,12,13,14,15,16,17,18,19,20} which is greater, the probability of choosing a multiple of 3 or a multiple of 4?[43]

> Two of the balls are red, 3 are gray and 5 are black.... What is greater, the probability of getting a red or a black ball? of getting a gray or a black ball?[44]

In this study a subject's understanding of this concept was determined by his performance on a test in which he was asked to compare the probabilities of winning a simple game in one trial under two different conditions. Half of the items represented situations in which the probabilities for success were equal.

Summary of the Problem

This study is a status study examining the performances of a random sample of 528 fourth, fifth, sixth and seventh grade children on three tests relating to the application of three basic concepts of probability to simple game situations. The subjects had not been taught probability in the schools before the administration of the tests.

---

[42] Flavell, op. cit., p. 346

[43] Duncan, et. al., op. cit., p. 255

[44] Ibid., p. 254

The subjects were divided into twenty-four groups on the basis of grade in school (four levels), I.Q. range (three levels) and sex (two levels). The results of the tests will be analyzed to determine if the mean performances of the groups differed significantly and if there are any significant differences due to interactions of the groups.

The tests consisted of a twelve item test on the concept of sample space, a twelve item test on the concept of probability of a simple event and a ten item test on the quantification of probabilities. Each item on the second test relates to a situation similar to the situation presented in the corresponding item on the first test. The first and second tests are also divided into two subtests of six items each, the first subtest involving only simple counting and the second subtest involving fundamental notions of combinations. A correlation study will be made between the scores on the three main tests and the subtests.

Hypotheses and Questions

One purpose of the study was to examine the relationship between the factors of I.Q., sex and grade and the subjects' performances on three probability tests. The specific hypotheses to be tested are:

1. There is no difference in the mean performances of children in the three I.Q. groups.

2. There is no difference in the mean performances of boys and girls.

3. There is no difference in the mean performances of children in the four grades.

4. There is no difference in the mean performances of children in the three I.Q. groups across the two sex groups.

5. There is no difference in the mean performances of children in the three I.Q. groups across the four grade levels.

6. There is no difference in the mean performances of boys and girls across the four grade levels.

7. There is no difference in the mean performances of children in the three I.Q. groups across the two sexes and four grade levels.

Another purpose of the study was to explore: 1) the relationship between language, non-language and total I.Q. data and the performance scores on the probability tests; 2) the relationship between the performance scores on the three main tests and the subtests. The specific questions to be asked for this part of the study are:

8. Which of the three available scores on the California Test of Mental Maturity; Language I.Q., Non-Language I.Q., or Total I.Q. is the best predictor of the performance scores on the three probability tests?

9. What is the relationship between the performance scores on the three probability tests within each grade?

10. What is the relationship between the performance scores on Subtest I-A and Subtest II-A within each grade?

11. What is the relationship between the performance scores on Subtest I-B and Subtest II-B within each grade?

Chapter III

PROCEDURES AND DESIGN OF THE STUDY

Population

The study was conducted in the Wausau, Wisconsin Public School System in November and December, 1967. The Wausau School District comprises an area of 256 square miles which includes the city of Wausau and the adjoining townships of Rib Mountain, Stettin, Berlin, Main and Texas. There are eighteen schools in the Wausau School District: one elementary school, grades K-4; fourteen elementary schools, grades K-6; two junior high schools, grades 7-9; and one high school, grades 10-12. The high school, junior high schools and ten of the elementary schools are in the city of Wausau with the other five elementary schools located in the adjoining townships. The total Public School enrollment in December, 1967 was approximately 9,500.

The immediate Wausau area is primarily industrial including more than eighty diversified manufacturing establishments. The adjoining townships include small residential communities as well as many small and large farms. The population from which the sample was drawn represented both the urban and rural areas of the district. The population of the district also represented a wide range of socio-economic backgrounds.

The population for the study consisted of 2,169 fourth, fifth, sixth and seventh grade children. This population represented approximately 87% of the total number of children enrolled in these grades in the district in October, 1967. The population included all children enrolled in grades four through seven for whom a Total I.Q. on the California Test of Mental Maturity [45] was available from the school files. The California Test of Mental Maturity is administered each year in the Wausau district as part of the regular testing program to all children enrolled in the third grade and the sixth grade. Therefore the I.Q. data for children in the population were obtained from tests administered during three different school years. Table 1 shows the total number of children enrolled in each grade, the number of children included in the population, the date of administration of the California Test of Mental Maturity, and the test form used for each grade level.

The frequencies of the Total I.Q. by grade and the frequencies for the total population are presented in Table 2. The population was partitioned into three groups, with approximately the same number of children in each group, on the basis of I.Q. The I.Q. range of 71-104 was selected as the first I.Q. range, 105-113 was selected as the second I.Q. range and 114-144 was selected as the third I.Q. range.

Table 3 gives the frequency of the grouped I.Q. for each grade

---

[45] Elizabeth T. Sullivan, Willis W. Clark, Ernest W. Tiegs, California Short-Form Test of Mental Maturity, 1963 Revision (Monterey, California: California Test Bureau, 1964)

Table 1

Population by Grade, Dates of Administration and Form of the California Test of Mental Maturity Used for Selection of the Population

Grade	Total Number of Children Enrolled	Number of Children in the Population	Date I.Q. Test Administered	Form of Test Used
4	632	603	October, 1966	Primary Short-Form S
5	635	526	October, 1965	Primary Short-Form S
6	566	459	October, 1964	Primary Short-Form S
7	672	581	January, 1967	Intermediate Short-Form S

and for the total population. Table 4 presents the mean I.Q. and standard deviation for each grade and for the total population.

It is clear from Table 3 that the distributions of I.Q.'s are not the same for all grade levels. The greatest discrepancy is apparent in grade seven where more than half of the population is in the third I.Q. range. This may be due to the fact that these children took a different form of the California Test of Mental Maturity at the sixth grade level.

A further restriction on the population for this study was that only children who had not received formal instruction on the topic of

Table 2

Frequencies of Total I.Q.'s for Each Grade and Total Population

I.Q.	Frequency				
	Fourth Grade	Fifth Grade	Sixth Grade	Seventh Grade	Total Population
71	1	-	1	3	5
72	-	-	1	-	1
73	-	-	-	-	-
74	1	-	-	-	1
75	-	-	-	1	1
76	-	-	1	1	2
77	-	-	-	2	2
78	-	-	-	-	-
79	-	-	2	1	3
80	-	1	2	2	5
81	1	-	1	1	3
82	-	-	3	3	6
83	-	1	1	-	2
84	-	3	-	1	4
85	2	1	5	3	11
86	2	4	1	5	12
87	1	-	3	1	5
88	4	-	4	7	15
89	-	1	4	3	8
90	1	-	2	4	7
91	4	4	4	3	15
92	3	2	4	2	11
93	9	5	3	2	19
94	14	5	5	2	26
95	9	9	11	13	42
96	13	10	10	4	37
97	11	10	7	3	31
98	20	16	16	8	60
99	27	17	8	3	55
100	22	17	15	4	58
101	14	24	11	13	62
102	22	7	14	14	57
103	21	17	19	12	69
104	24	19	11	11	65
105	35	27	12	13	87

Table 2 (continued)

I.Q.	Fourth Grade	Fifth Grade	Sixth Grade	Seventh Grade	Total Population
106	16	19	22	9	66
107	22	21	16	9	68
108	25	22	17	22	86
109	31	37	12	20	101
110	25	23	19	20	87
111	12	17	17	14	60
112	23	18	9	15	65
113	22	25	12	30	89
114	18	20	10	24	72
115	16	13	8	22	59
116	24	22	11	17	74
117	19	9	13	15	56
118	14	17	8	30	69
119	14	21	10	16	61
120	8	6	9	23	46
121	5	8	13	22	48
122	12	6	10	25	53
123	12	5	3	20	40
124	6	6	7	22	41
125	7	2	7	17	33
126	5	2	6	11	24
127	2	1	4	6	13
128	1	2	4	10	17
129	2	3	1	5	11
130	1	2	6	3	12
131	1	1	8	3	13
132	-	-	3	2	5
133	-	-	-	1	1
134	-	-	2	2	4
135	-	-	2	1	3
136	-	-	1	-	1
137	-	-	1	-	1
138	-	-	2	-	2
139	-	-	1	-	1
140	-	-	2	-	2
141	-	-	-	-	-
142	-	-	-	-	-
143	-	-	-	-	-
144	-	-	1	-	1

Table 3

Frequencies of Grouped Total I.Q.'s for the Population at Each Grade Level and the Total Population

I.Q. Range	Fourth Grade	Fifth Grade	Sixth Grade	Seventh Grade	Total Population
$71 \leq$ I.Q. ≤ 104	226	171	169	133	699
$105 \leq$ I.Q. ≤ 113	211	209	137	151	708
$114 \leq$ I.Q. ≤ 144	166	146	153	297	762
Totals	603	526	459	581	2169

Table 4

Mean Scores and Standard Deviations of Total I.Q.'s for the Population at Each Grade Level and Total Population

	Fourth Grade	Fifth Grade	Sixth Grade	Seventh Grade	Total Population
Mean I.Q.	107.84	108.58	108.80	111.81	109.28
Standard Deviation	8.12	5.73	12.46	12.55	10.19

probability would be considered. The aritmetic textbook series[46] used in all intermediate grades and the mathematics textbook[47] used in the seventh grade do not contain any units on probability so this topic was not included in the regular mathematics curriculum for pupils in the Wausau district. As a further check all classroom teachers, grades four through six, and all seventh grade mathematics teachers were polled to see if they had taught any ideas of probability as supplementary units. All but two of the teachers replied that they had taught no probability units in their arithmetic classes. One fifth grade teacher indicated that she had spend part of one class period (approximately 15 minutes) discussing a coin tossing experiment. This discussion was very informal and did not introduce probability terminology. It was decided that these children had not received enough training from this one very brief session to bias the study so this class was not eliminated from the population. Another teacher reported that one of his sixth grade children was working on a unit on probability as an independent project. This child was not included in the population. All elementary principals in the district indicated that, to the best of their knowledge, none of the fourth, fifth and sixth grade teachers who had left the system had taught probability

[46] E. T. McSwain, Kenneth E. Brown, Bernard H. Gundlach, Ralph J. Cooke, Arithmetic 4, 5 and 6 (River Forest, Illinois: Laidlow Brothers, 1965)

[47] Henry Van Engen, Maurice Hartung, Harold Trimble, Emil Berger, Ray Cleveland, Seeing Through Mathematics, Books 1 and 2 (Chicago: Scott, Foresman and Company, 1961)

in their arithmetic classes. Therefore it is reasonable to assume that the 2,169 children in the population had not had any formal learning experiences with probability before the tests for this study were administered.

Subjects

The population of 2,169 fourth, fifth, sixth and seventh grade children was partitioned into twenty-four subclasses on the basis of grade, sex and I.Q. range. The number of children in each of the twenty-four subclasses is given in Table 5.

TABLE 5

Number of Children in Each of the Twenty-four Subclasses of the Total Population

I.Q. Range	Sex	Grade 4	Grade 5	Grade 6	Grade 7
Range I 71-104	Male	122	86	86	64
	Female	104	85	83	69
Range II 105-113	Male	120	118	80	62
	Female	91	91	67	89
Range III 114-144	Male	80	75	85	136
	Female	86	71	68	161

A random sample of twenty-two children was selected from each of the twenty-four subclasses shown in Table 5 for a total sample of 528 children. Schools were not considered as a variable so no attempt was made to have each school represented in each of the subclasses. The total sample did include some children from each of the fifteen elementary schools and each junior high school.

After the initial random selection of twenty-two subjects for each cell, the random selection process was continued so that each elementary school had one alternate, who was enrolled in that school, for each subclass represented in that school. Since all of the six subclasses in the seventh grade were represented in each junior high school, three alternates from each school were selected for each of these six subclasses. As would be expected several of the alternates were used in place of children who were absent on the day the tests were administered. The number of alternates selected was adequate and it was not necessary to schedule any special test sessions in any of the schools.

The Pilot Studies

Several pilot studies were conducted to help answer the following questions:

1. Can the tests be administered to a large group of subjects?

2. Is there any evidence to indicate that young children do possess some understanding of the concepts of probability under investigation?

3. Is there any evidence to indicate that some relationship exists between I.Q., grade, arithmetic achievement and an understanding of the probability concepts under investigation?

4. Is there any evidence of reliability among items on the three probability tests?

In addition the pilot studies were designed to help decide: what probability terminology would be most appropriate; the type of sample items to include in the instructions so the subjects would understand the questions to be asked; which items were ambiguous or misleading; and whether or not the reading level of the items was too high.

In the first pilot study, conducted in October, 1966, three 8 item tests were administered to 103 subjects in an elementary school in Monona, Wisconsin. The sample consisted of one sixth grade classroom (27 subjects), one fifth grade classroom (27 subjects) and two fourth grade classrooms (49 subjects). The subjects in each grade represented a wide range of general intelligence and arithmetic achievement. The tests were administered to classroom groups of from 23-27 subjects.

The results of the first pilot study provided affirmative information for questions 1, 2, 3 and 4. In addition this study indicated the relative difficulty of the test items and suggested necessary changes in the wording and presentation of the items. The study also indicated that it may be appropriate to include seventh grade subjects in the experiment.

The tests were revised and second pilot study was conducted in December, 1966. The sample for this study included 28 seventh grade, 53 sixth grade, 49 fifth grade, and 55 fourth grade subjects from a different elementary school in Monona, Wisconsin. The results of the second study were consistent with the results of the first study giving affirmative information for questions 1-4.

The tests were revised again using the item analysis data and other information gained from the second pilot study. A third pilot study was conducted in October, 1967. The sample for this study consisted of 60 sixth grade and 43 fourth grade subjects from an elementary school in Madison, Wisconsin. The information gained from the third pilot study helped determine the final form and content of the tests used in the study.

The Tests

The purpose of this study is to assess the children's intuitive understanding of three concepts of probability. To achieve this three tests were constructed,[48] one for each concept, with items for which the child's responses would indicate if the child could apply the concepts in a variety of simple experiment and game situations.

The words used in constructing the test items were considered to be words or blends of words that would be familiar to fourth grade children. A check of the test items by a reading consultant in the Wausau schools confirmed the writer's conclusion that the items would not present any unusual reading difficulties for the subjects in the study. The meaning of each of the underlined words in the test items was explained, in terms of the context of the situation in which it was being used, in the preliminary instructions to the children.

The items in each test were arranged in what was considered to be

[48]See Appendix A for the test items.

an order of difficulty, from easiest to most difficult. This order was determined from the results of the pilot studies.

The first test consisted of twelve items on the concept of sample space. Each item described a lot-drawing experiment, defined an outcome of the experiment and included a diagram which illustrated the objects used in the experiment. In each item where an outcome was a pair of things, a sample of how this outcome could be represented was included in the item. The first four items involved only simple counting. Items five and six involved sampling without replacement. The last six items involved relatively simple ideas of combinations. Item seven involved the number of combinations of three things taken two at a time. Item eight involved the number of combinations of four things taken two at a time. Item nine involved the number of combinations of four things taken three at a time. Item ten involved the number of pairs in a 2 x 2 cartesian product space. Item eleven involved the number of pairs in a 2 x 3 cartesian product space. Item twelve involved the number of pairs in a 3 x 3 cartesian product space. In items ten through twelve the pairs were considered as combinations, ignoring the order of elements in each pair.

The response for an item in Test I was to list all of the different outcomes possible for the experiment described in the item. The items were scored either right or wrong. An item was scored right if all of the different outcomes possible for the experiment were listed. The outcomes could be listed in any order and all outcomes involving pairs of things were considered as combinations or unordered pairs. An item was scored wrong if the child's list did not include all of the different outcomes or if the list contained more outcomes than were possible for the particular experiment described in the item.

The second test consisted of twelve items on the concept of probability of a simple event. Each item in the second test presented a lot-drawing situation similar to the situation in the corresponding item on the first test. The situations in the items of Test II were described in terms of a simple game situation rather than as an experiment as in the items on Test I. An implicit question in each item in the second test is: How many different outcomes are possible on each draw? This is the explicit question asked in the corresponding item on the first test.

In each item on Test II the rules for playing the lot-drawing game were described. The way in which the game could be won was clearly specified. The ways in which one would lose the game were described in a general manner. In addition, just as in Test I, a diagram illustrating the objects used to play the game was included with each item.

The response for an item in Test II was to fill in the blanks in the expression _____ out of _____ which indicated the chance of winning the game described in the item if one was allowed only one draw; i.e., only one opportunity to play the game. This expression represented the child's interpretation of the probability of the simple event described in the item as "the way to win the game." The items were scored either right or wrong. An item was scored right if the child's response represented the correct probability for the event described in the item.

For example, in item 14, "2 out of 12" and "1 out of 6" were both considered correct responses for the item. All other responses which did not represent the correct probability, including statements representing "odds of winning" were considered wrong.

The third test consisted of ten multiple-choice items on the concept of quantification of probabilities. Each item described a lot-drawing game. The way in which one would win the game was clearly specified. Each item also included a general description of the ways one would lose the game. The diagram accompanying each item consisted of pictures of two boxes (or spinners) which illustrated the objects that could be used to play the game.

The response for an item on Test III was to select one of three possible choices, "Box A," "Box B," or "It doesn't make any difference." The response represented the child's interpretation of the better probability of success for a simple event in one trial under two different conditions. A selection of the third choice indicated that the child considered the event to have the same probability of success in each situation.

Items 26, 30, 31, 32 and 33 represented situations in which the specified simple event had the same probability of success under both conditions.

The items on Test III were scored either right or wrong.

The three tests were put together in the form of a booklet with two items on each page. All items were numbered consecutively from 1 - 34. A blank blue sheet was inserted after item 12 to separate Test I and

Test II. A blank yellow sheet was inserted after item 24 to separate Test II and Test III.

Administering the Tests

The tests were administered during the last week of November and first two weeks of December, 1967. Tests were administered to children in all of the fifteen elementary schools and two junior high schools in the district. A test schedule was arranged with the cooperation of the school administrators so that all testing in a particular school could be done on that same day. The tests were administered to groups of subjects, and wherever possible all subjects in a particular school were tested at the same time. The test groups in the elementary schools varied from 8 to 27 children. In three of the elementary schools the children were tested in two separate groups at different times during the same day. The groups were split in these schools so that the group size would be less than 30 for each test session. The same tests were administered to all children in the sample so the elementary school groups were mixed, including children from the fourth, fifth and sixth grades. Seventh grade subjects were tested separately in the junior high schools. Because scheduling problems for the junior high schools were more complex and space for administering group tests was more difficult to obtain, the seventh grade subjects were tested in groups that were larger than the elementary school groups. In one junior high school the children were tested in two groups with approximately 32 children in each group. In the second junior high school the children were tested in one large group of 68 children. The large groups at the seventh grade level did not create any unusual difficulties either in giving instructions for the tests or in proctoring the tests.

The tests were administered in a variety of rooms depending on the facilities available in the individual schools. Precautions were taken so that subjects would not be able to see their neightbor's paper. The children were seated in alternate desks in a classroom or at widely spaced tables in a library, cafeteria or gymnasium. Cardboard dividers, approximately 30 inches high, were used on the tables to help isolate the children.

All tests were administered by the writer. The same procedure was followed for each test session. A brief 5-8 minute warm-up period preceded each session during which time the writer introduced himself, checked the roll to see if all children scheduled for that time were present and sent for alternates when necessary. This time was also used to explain the purpose of the test session. The children were assured that the set of questions they were being asked to answer would not be considered a school test; no grades would be given; results would not be shown to either their teachers or parents; and the answers they gave would in no way affect the evaluation they would receive for their regular school work. The children were encouraged to do their best work and not to be concerned if they were uncertain about answers for some of the questions. They were also encouraged to answer all of the questions on each test and told to write their best "guess" if they were not certain about an answer. Pencils were distributed to children who had not brought their own. Some time was allowed for general questions.

The test booklets were then distributed. The children were instructed to print their name, grade and school on the cover of the booklet in the space provided but not to open the booklet until told to do so.

The instructions[49] for Test I were read to the children so that all groups received the same instructions. A box and colored cards were used to demonstrate the sample items for this test. The sample items and correct responses for the sample items were written on a chalkboard or large white cardboard so that they were clearly visable by the entire group.

After the presentation of the sample items the children were asked if there were any questions. They were also reminded that questions would not be answered after the test was begun. After all questions were answered the children were asked to open their test booklets to page 1. The children were told not to work ahead but to wait for instructions for each item. The first question was read aloud by the writer while the subjects read the item silently. The children were then asked to write their answer for this item and to cover their answer with their scratch paper when finished. After all of the children had written their answers for the first item, the same process was repeated for item 2. The items were read aloud to eliminate reading difficulties

[49] See Appendix A for the instructions for the tests.

as much as possible, to emphasize underlined words in the items, and to emphasize the importance of using the diagram in the item to help answer the question. There was no time limit for the items and children were asked to be patient after answering an item to allow everyone enough time to answer the question before proceeding to the next item. The same process was repeated for items 3, 4, 5 and 6. After item 7 was read aloud the children were told to continue working on their own through item 12 and to stop when they reached the blank blue sheet in their test booklet. It was very easy to tell when all children in the group had completed Test 1 by checking to see that all test booklets were open to the blue page.

After all children in the group had completed Test I the instructions for Test II were read aloud to the group. The sample item was written on a chalkboard or large white card. A box and colored cards were used to demonstrate the sample item. The response for the sample item was written under the item. A brief explanation was given to show why the correct response for the sample item was 1 out of 4 rather than 1 out of 3. This explanation included a reminder that, in order to answer a question like the sample item, the child should consider two important questions: 1) How many different things can happen when playing the game? 2) How many ways can you win the game?

The children were then told to open their booklets to page 7, item 13. This item was read aloud by the writer while the subjects read it silently. It was pointed out that the question was to be answered by filling in the blanks _____ out of _____. The children were then told to answer item 13 and to then work ahead on their own through item 24.

They were told to stop and wait for the rest of the group when they reached the blank yellow sheet in their test booklet.

After all the children in the group had finished Test II the instructions for Test III were read aloud to the group. No sample items were presented for Test III. The children were instructed to work on the items at their own speed. Subjects were excused to return to their classroom after they had completed all of the items on Test III.

The time required to administer the three tests, including the warm-up demonstrations, varied from 40-60 minutes.

Design of the Study

The basic design for the study was a 3 x 2 x 4 multivariate analysis of covariance design with 22 subjects in each cell. The three factors (I.Q.: Range I, 71-104; Range II, 105-113; Range III, 114-144; Sex: male, female; Grade: fourth, fifth, sixth, seventh) were assumed to be fixed and completely crossed. The covariates were grade equivalent scores on the three parts of the Stanford Arithmetic Achievement Test: computation, concepts and applications.[50] The dependent variables were the performance scores of the three probability concept tests described in a previous section in this chapter. A schemetic representation of the design is given in Table 6.

Since multivariate analysis of covariance was used as the basic design for the study, the mean vectors in the twenty-four cells of

---

[50]T. L. Kelley, R. Madden, E. F. Gardner, H. C. Rudman, Stanford Achievement Test (New York: Harcourt, Brace and World, Inc., 1964)

design represent results obtained from scores that have been adjusted for the covariates.

Table 6

Schematic Representation of the Design of the Study

I.Q.	Sex	Grades C_1	C_2	C_3	C_4
A_1	B_1	$\underline{X}_{111.}$	$\underline{X}_{112.}$	$\underline{X}_{113.}$	$\underline{X}_{114.}$
	B_2	$\underline{X}_{121.}$	$\underline{X}_{122.}$	$\underline{X}_{123.}$	$\underline{X}_{124.}$
A_2	B_1	$\underline{X}_{211.}$	$\underline{X}_{212.}$	$\underline{X}_{213.}$	$\underline{X}_{214.}$
	B_2	$\underline{X}_{221.}$	$\underline{X}_{222.}$	$\underline{X}_{223.}$	$\underline{X}_{224.}$
A_3	B_1	$\underline{X}_{311.}$	$\underline{X}_{312.}$	$\underline{X}_{313.}$	$\underline{X}_{314.}$
	B_2	$\underline{X}_{321.}$	$\underline{X}_{322.}$	$\underline{X}_{323.}$	$\underline{X}_{324.}$

Key

A_1 = Range I, 71-104	B_1 = male	C_1 = fourth grade
A_2 = Range II, 105-113	B_2 = female	C_2 = fifth grade
A_3 = Range III, 114-144		C_3 = sixth grade
		C_4 = seventh grade

All underlined symbols represent mean vectors of order 3 x 1.

Each cell entry is a vector $\underline{X}_{ijk} = \frac{\sum_{m=1}^{n_{ijk}} \underline{X}_{ijkm}}{n_{ijk}}$, which represents a vector of order p x 1, the mean vector of n_{ijk} observation vectors, one for each subject.

The subscripts i, j and k represent the levels of the factors, p represents the number of dependent variables and n_{ijk} represents the number of subjects in each cell. For this study, i = 1, 2, 3; j = 1, 2; k = 1, 2, 3, 4; p = 3; and $n_{ijk} = 22$ for all i, j, k.

For purposes of analysis it was assumed that:

$E(\underline{X}_{ijkm}) = \underline{\mu} + \underline{\alpha}_i + \underline{\beta}_j + \underline{\xi}_k + \underline{\alpha\beta}_{ij} + \underline{\alpha\xi}_{ik} + \underline{\beta\xi}_{jk} + \underline{\alpha\beta\xi}_{ijk}$. That is, the expected value is a linear function of the main effects and the interaction effects.

Jones[51] reports that for a design with an equal number of observations in each cell, the computation and decomposition of total sums of products of a multivariate analysis of variance design is directly analogous to that of sums of squares in univariate analysis of variance. Winer [52] outlines a design for a three-factor analysis of

---

[51]Lyle V. Jones, Some Illustrations of Psychological Experiments Designed for Multivariate Statistical Analysis. Report Number 28, The Psychometric Laboratory (Chapel Hill, North Carolina: University of North Carolina, December, 1960) p. 3.

[52]B. J. Winer, Statistical Principals in Experimental Design (New York: McGraw-Hill Book Company, 1962), pp. 170-171.

variance. Using the models presented by Winer and the two-factor multivariate model presented by Jones[53] the multivariate analysis of variance (MANOVA) table for this study was constructed and is represented by Table 7.

This design will be used to Test Hypotheses 1-7. Pearson product moment correlation coefficients will be calculated to answer Questions 8-11 of the statement of the problem.

---

[53]Jones, op. cit., p. 4

Table 7

MANOVA Table

Source	Degrees of Freedom	Sum of Products	Mean Product
A	$n_{h_A}=a-1$	$S_A=bcn\sum_{i=1}^{a}\overline{A}_i-\overline{G}$	$M_A=\frac{S_A}{a-1}$
B	$n_{h_B}=b-1$	$S_B=acn\sum_{j=1}^{b}\overline{B}_j-\overline{G}$	$M_B=\frac{S_B}{b-1}$
C	$n_{h_C}=c-1$	$S_C=abn\sum_{k=1}^{c}\overline{C}_k-\overline{G}$	$M_C=\frac{S_C}{c-1}$
AxB	$n_{h_{AB}}=(a-1)(b-1)$	$S_{AB}=cn\sum_{i=1}^{a}\sum_{j=1}^{b}(\overline{AB}_{ij}-\overline{A}_i-\overline{B}_j)+\overline{G}$	$M_{AB}=\frac{S_{AB}}{(a-1)(b-1)}$
AxC	$n_{h_{AC}}=(a-1)(c-1)$	$S_{AC}=bn\sum_{i=1}^{a}\sum_{k=1}^{c}(\overline{AC}_{ik}-\overline{A}_i-\overline{C}_k)+\overline{G}$	$M_{AC}=\frac{S_{AC}}{(a-1)(c-1)}$
BxC	$n_{h_{BC}}=(b-1)(c-1)$	$S_{BC}=an\sum_{j=1}^{b}\sum_{k=1}^{c}(\overline{BC}_{jk}-\overline{B}_j-\overline{C}_k)+\overline{G}$	$M_{BC}=\frac{S_{BC}}{(b-1)(c-1)}$

Table 7 Continued

Source	Degrees of Freedom	Sum of Products	Mean Product
AxBxC	$n_{h_{ABC}}=(a-1)(b-1)(c-1)$	$S_{ABC}=n\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}(\overline{ABC}_{ijk}-\overline{AB}_{ij}-\overline{AC}_{ik}-\overline{BC}_{jk}+\overline{A}_i+\overline{B}_j+\overline{C}_k)-\overline{G}$	$M_{ABC}=\frac{S_{ABC}}{(a-1)(b-1)(c-1)}$
Within cells	$n_e=abc(n-1)$	$S_e=S_T-S_{ABC}-S_{BC}-S_{AC}-S_{AB}-S_C-S_B-S_A$	$M_e=\frac{S_e}{abc(n-1)}$
Total	abcn-1	$S_T=\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{m=1}^{n}\overline{X}_{ijkm}-\overline{G}$	

Key

$\overline{A}_i=\underline{X}_{i\ldots}\underline{X}'_{i\ldots}$	$\overline{AB}_{ij}=\underline{X}_{ij..}\underline{X}'_{ij..}$	$\overline{ABC}_{ijk}=\underline{X}_{ijk.}\underline{X}'_{ijk.}$
$\overline{B}_j=\underline{X}_{.j..}\underline{X}'_{.j..}$	$\overline{AC}_{ik}=\underline{X}_{i.k.}\underline{X}'_{i.k.}$	$\overline{X}_{ijkm}=\underline{X}_{ijkm}\underline{X}'_{ijkm}$
$\overline{C}_k=\underline{X}_{..k.}\underline{X}'_{..k.}$	$\overline{BC}_{jk}=\underline{X}_{.jk.}\underline{X}'_{.jk.}$	$\overline{G}=abcn\underline{X}_{\ldots.}\underline{X}'_{\ldots.}$

Each underlined symbol represents a mean vector of order p x 1.

Each overlined symbol represents a mean product matrix of order p x p.

Chapter IV

RESULTS OF THE RELIABILITY STUDIES

Description of the Statistics Used in the Reliability Studies

Reliability studies were carried out for each of the three probability concept tests and four subtests.[54] A separate reliability study was carried out for each test and subtest based on the performance of the 132 subjects at a particular grade level. A total of twenty-eight reliability studies were conducted, seven for each of the four grade levels included in this study. The results of these reliability studies are reported in this chapter.

For each reliability study the following information is reported: 1) a frequency distribution of the total scores for the test or subtest under consideration; 2) a two-way ANOVA table for the individuals and items under consideration and a Hoyt reliability coefficient of internal consistency; and 3) an item analysis which includes the following statistics for each item: a difficulty index; item-criterion correlation; X_{50}; and β.

The Hoyt reliability coefficient (H.R.) is a measure of internal

---

[54]Frank B. Baker, Test Analysis Package: A Program for the CDC 1604-3600 Computers, (Madison: University of Wisconsin, Laboratory of Experimental Design, Department of Educational Psychology, June, 1966).

consistency among test items. This coefficient is given by the formula[55]

$$r_h = \frac{MS_{ind} - MS_{ind \times items}}{MS_{ind}} = 1 - \frac{MS_{ind \times items}}{MS_{ind}}$$

where MS_{ind} and $MS_{ind \times items}$ are two mean squares in a two-way analysis of variance in which one of the factors in the subjects and the other is the test items.

The purpose of an item analysis is to describe how each item on the test functions. The four statistics used to describe the items of the tests used in this study are based on the item characteristic curve, which is considered to be a fundamental concept in item analysis. According to Baker, "An item characteristic curve is a smooth curve fitted to the proportion of persons at each criterion score level who made the particular response being studied."[56] He goes on to say, "If one assumes the item characteristic curve has the form of a normal ogive the parameter of the normal curve (μ, σ) can be used to describe the data."[57] Assuming that the item characteristic curve has the form of the normal ogive Baker gives the following definitions for X_{50} and β.

> The parameters of the item characteristic curve which specify the normal ogive fitted to the item response data are the following:

---

[55]Frank B. Baker, Empirical Determination of Sampling Distributions of Item Discrimination Indices and a Reliability Coefficient (Department of Educational Psychology, School of Education, University of Wisconsin, Contract E-2-10-071, November, 1962), p. 87.

[56]Frank B. Baker, "An Intersection of Test Score Interpretation and Item Analysis, Journal of Educational Measurement, I (June, 1964), p. 24.

> X_{50}, the criterion score at which the probability of correct response is .5. The parameter is expressed in units of the criterion variable standard deviation.
>
> β a measure of the steepness of the item characteristic curve which specifies the capability of the item to discriminate between the individuals possessing various amounts of the criterion ability. This parameter is the reciprocal of the standard deviation of the fitted normal ogive.[58]

The discrimination index β can also be interpreted in a less-technical way as "the slope of the item characteristic curve at X_{50}."[59] Even though this interpretation is not mathematically correct it provides a usable approximation for practical purposes.

The index of item difficulty reported in this study represents the proportion of the total group who gave the correct response for the item. However, item difficulty is also related to the item indices X_{50} and β. "Item difficulty corresponds to the area under the item characteristic curve and is hence a function of X_{50} and β."[60]

Biserial correlation may be used to compute the correlation between a criterion variable, such as a total test score, and an item response variable if it is assumed that the response variable is continuous and normally distributed but obtainable only as a dichotomus response.[61] The formula for computing the biserial correlation coefficient is[62]

---

[58]Baker, Empirical Determination of ..., pp. 11-12

[59]Baker, An Intersection of ..., p. 25.

[60]Baker, Empirical Determination of ..., p. 29.

[61]Ibid., p. 5

[62]Ibid., p. 6

$$r_b = \frac{\overline{X}_1 - \overline{X}}{S_x} \left(\frac{P}{z}\right)$$

where: $\overline{X}_1$ is the mean score of all persons answering the item correctly

$\overline{X}$ is the mean of the sample

S_x is the standard deviation of the sample

P is the proportion of persons answering the item correctly

z is the ordinate of the normal curve at the deviate which divides the area of the unit normal curve into P and (1-P)

The computational formula $\beta = \frac{r_b}{\sqrt{1 - r_b}}$ expresses a relationship between r_b and β and is used in the computer program to obtain β.[63]

In all of the reliability studies carried out for this study the total score on the items under consideration was used as the internal criterion measure for computing the item-analysis statistics.

As previously noted reliability studies for each test and subtest were conducted individually for each of the grades; four, five, six and seven. Thus each reliability study is based on the performances of the 132 subjects in the sample for a particular grade. In the reports of the reliability studies that follow, some of the results for all four grades are presented in the same table so that the results can be compared more easily.

Reliability Studies of Test I (Sample Space)

Test I is a twelve item test[64] on the concept of points of a finite sample space. Each item was scored either right or wrong.

---

[63]Baker, Test Analysis Package:..., p. 6.

[64]See Appendix A for the test items. Test I consists of items 1-12.

Therefore the range of total scores possible for Test I was from 0 to 12 inclusive.

Tables 8 and 9 present the frequency distributions, mean scores and standard deviations for the scores on Test I for all grades. Table 8 shows that the distributions of scores vary for the different grades but there is some similarity between the distributions for grades six and seven. The differences between grades is also reflected by the mean scores reported in Table 9. The total scores ranged from 0 to 12 for grades four, five and seven and from 1 to 12 for grade six. This wide range of scores within each grade indicates that the performances of children on Test I varied considerably. An inspection of Tables 10-13 shows that the difference between individuals at each grade is significant at the 1% level. A more detailed analysis of differences in performances of individuals will be given in Chapter V.

Tables 10-13 give the ANOVA tables and Hoyt reliability coefficients computed for Test I. The reliabilities for grades four, five, six and seven are .81, .81, .76 and .81 respectively.

Table 14 contains the item analysis of Test I for each grade. The item statistics included in the table were computed using the total test score as the criterion variable. The item statistics included in the table are: the item difficulty, r_b, X_{50} and β.

All of the items were relatively easy for grades six and seven. The item difficulties for these grades range from .49 to .94 with the majority being above .60. Items 1-6, which involve only simple counting, were relatively easy for grades four and five also. Only item 6, grade four, has a difficulty index less than .50. As would be expected items 7-12, which involve combinations are more difficult, particularly for grade four. The item difficulties for items 8-12 range from .24 to .31 for grade four. A more detailed analysis of the types of errors made on these items is given in Chapter V.

The item criterion correlations are generally high. The exception is the correlation for item 1, grade six, which is only .31. However, only eight pupils in grade six answered this item incorrectly. Since this was the first item on the test the errors may have been random which would account for the low correlation. Item 11, grade seven, has a correlation greater than 1. The assumption on which the computation of a biserial correlation is based is that the variable underlying the dichotomy is continuous and normal. Ferguson explains that if this assumption is violated irregularities can occur.

> Theoretically, the maximum and minimum values of r_{bi} are independent of the point of dichotomy and are -1 and +1. An implicit assumption underlying this statistic is that the continuous many-valued variable is normal, as well as the variable underlying the dichotomy. Values of r_{bi} greater than unity can occur under gross departures from normality.[65]

Whenever a value of r_b greater than unity is obtained, β is set equal to zero by the GITAP computer program used in this study.

The items appear to be good discriminators at all grade levels. The majority of β's are .70 or greater. A β of .70 indicates that the slope of the item characteristic curve at the X_{50} point is approximately 35 degrees.

---

[65]George A. Ferguson, Statistical Analysis in Psychology and Education (New York: McGraw-Hill Book Company, Inc., 1959), pp. 203-204.

Table 8

Frequency Distributions of Total Scores on Test I for Grades Four, Five, Six and Seven

Total Score	Frequency Grade 4	Grade 5	Grade 6	Grade 7
12	6	18	15	22
11	7	8	28	31
10	4	12	17	21
9	7	11	16	13
8	8	12	11	12
7	20	17	18	9
6	16	11	5	6
5	10	10	8	7
4	9	10	8	2
3	16	11	2	4
2	16	10	3	3
1	8	1	1	0
0	5	1	0	2

Table 9

Mean Scores and Standard Deviations for Test I

Grade	Mean Score	Standard Deviation
4	5.47	3.17
5	7.16	3.24
6	8.57	2.72
7	9.00	2.83

Table 10

ANOVA and Reliability Table for Test I, Grade 4

Source of Variation	d.f.	MS	F	R
Individuals	131	.8410	5.1560**	0.81
Items	11	4.5591	27.9515**	
Error	1441	.1631		
Total	1583			

** $p < .01$

Table 11

ANOVA and Reliability Table for Test I, Grade 5

Source of Variation	d.f.	MS	F	R
Individuals	131	.8836	5.3905**	0.81
Items	11	2.6799	16.3481**	
Error	1441	.1639		
Total	1583			

** $p < .01$

Table 12

ANOVA and Reliability Table for Test I, Grade 6

Source of Variation	d.f.	MS	F	R
Individuals	131	.6224	4.1045**	0.76
Items	11	2.1283	14.0355**	
Error	1441	.1516		
Total	1583			

** $p < .01$

Table 13

ANOVA and Reliability Table for Test I, Grade 7

Source of Variation	d.f.	MS	F	R
Individuals	131	.6743	5.1432**	0.81
Items	11	1.7948	13.6894**	
Error	1441	.1311		
Total	1583			

** $p < .01$

Table 14

Item Analysis for Test I for Grades Four, Five, Six and Seven

Item	Grade	Item Difficulty	r_b	X_{50}	β
1	4	.67	.54	- .81	.64
	5	.81	.73	-1.21	1.06
	6	.94	.31	-4.94	.33
	7	.92	.69	-2.07	.96
2	4	.71	.69	- .81	.94
	5	.71	.64	- .88	.83
	6	.72	.65	- .89	.86
	7	.80	.77	-1.07	1.21
3	4	.68	.69	- .66	.96
	5	.71	.62	- .87	.79
	6	.73	.58	-1.09	.70
	7	.79	.87	- .91	1.80
4	4	.58	.58	- .33	.71
	5	.79	.72	-1.12	1.03
	6	.86	.71	-1.53	1.02
	7	.89	.67	-1.78	.93
5	4	.50	.66	- .01	.89
	5	.65	.51	- .73	.60
	6	.81	.72	-1.22	1.04
	7	.83	.75	-1.23	1.16
6	4	.42	.81	.26	1.39
	5	.53	.68	- .10	.93
	6	.70	.81	- .66	1.36
	7	.78	.63	-1.22	.81
7	4	.56	.77	- .20	1.19
	5	.67	.64	- .67	.84
	6	.81	.65	-1.35	.85
	7	.77	.81	- .93	1.36
8	4	.31	.83	.60	1.48
	5	.49	.82	.03	1.42
	6	.67	.74	- .61	1.09
	7	.68	.63	- .75	.81

Table 14 (continued)

Item	Grade	Item Difficulty	r_b	X_{50}	β
9	4	.26	.81	.80	1.36
	5	.41	.78	.30	1.24
	6	.51	.48	- .04	.55
	7	.49	.68	.03	.94
10	4	.29	.81	.69	1.36
	5	.45	.92	.13	2.38
	6	.56	.80	- .19	1.35
	7	.62	.86	- .35	1.71
11	4	.24	.86	.80	1.67
	5	.47	.87	.07	1.80
	6	.62	.91	- .34	2.19
	7	.72	1.04	- .56	0
12	4	.26	.85	.75	1.64
	5	.47	.93	.07	2.57
	6	.62	.92	- .33	2.30
	7	.71	.96	- .58	3.55

Reliabilities Studies of Test II (Probability of a Simple Event)

Test II is a twelve item test[66] on the concept of the probability of a simple event in a finite sample space. Each item was scored either right or wrong. Therefore the range of total scores possible for Test II was from 0 to 12 inclusive.

Table 15 gives the frequency distributions of total scores for Test II. Table 16 contains the mean scores and standard deviations. Comparing these tables with Tables 8 and 9 shows that Test II was considerably more difficult than Test I for all grades. An inspection of Table 15 shows that the majority of scores lie in the range 0 to 8. More than 90% of the pupils in grades four and five and more than 70% of the pupils in grades six and seven had a total score of 6 or less. It is interesting to note that only one pupil answered all items correctly and this was a fifth grade pupil.

The reliability coefficients for Test II and analysis of variance used to compute these coefficients are reported in Tables 17-20. The coefficients are .62, .72, .72 and .73 for grades four, five, six and seven respectively. The reliabilities on Test II are slightly smaller than the reliabilities on Test I. This is probably due to the fact that seven of the items on Test II were very difficult items for all grades. Tables 17-20 also indicate that the differences among items and differences among individuals were highly significant for all grades.

---

[66] See Appendix A for the test items. Test II consists of items 13-24

Table 21 reports the item analysis for Test II. An examination of the item difficulties reported in this table shows that items 18-24 were difficult items for all grades. The item difficulties for items 20, 23 and 24 are .10 or less indicating that these items were extremely difficult for all pupils. These difficulties are not too surprising since item 18 involved obtaining the probability of a simple event when sampling without replacement and items 19-24 involved ideas of combinations. The X_{50} values reflect the relative difficulties of the items. Almost all of the X_{50}'s for items 18-24 are high positive values well above the mean of the criterion scores. Items 13-17 were relatively easy items, except for grade four.

Items 13-18, 23 and 24 are good discriminators with the majority of the β's greater than 1. Items 19-22 do not correlate very highly with the criterion scores at some of the grade levels and therefore have an erratic discrimination pattern. A more detailed analysis of the types of errors pupils made on the items in Test II is contained in Chapter V.

Table 15

Frequency Distributions of Total Scores on Test II for Grades Four, Five, Six and Seven

Total Score	Frequency Grade 4	Grade 5	Grade 6	Grade 7
12	-	1	-	-
11	-	1	2	5
10	-	1	-	3
9	-	1	2	2
8	2	2	13	10
7	6	4	16	19
6	7	9	16	20
5	12	16	13	23
4	24	24	17	16
3	21	18	24	11
2	20	27	15	13
1	24	13	9	6
0	16	15	5	4

Table 16

Mean Scores and Standard Deviations for Test II

Grade	Mean Score	Standard Deviation
4	2.90	2.02
5	3.31	2.33
6	4.51	2.44
7	5.15	2.53

Table 17

ANOVA and Reliability Table for Test II, Grade 4

Source of Variation	d.f.	MS	F	R
Individuals	131	.3433	2.6521**	0.62
Items	11	5.3519	41.3412	
Error	1441	.1295		
Total	1583			

** $p < .01$

Table 18

ANOVA and Reliability Table for Test II, Grade 5

Source of Variation	d.f.	MS	F	R
Individuals	131	.4544	3.5461**	0.72
Items	11	6.5709	51.2824**	
Error	1441	.1281		
Total	1583			

** $p < .01$

Table 19

ANOVA and Reliability Table for Test II, Grade 6

Source of Variation	d.f.	MS	F	R
Individuals	131	.5006	3.5767**	0.72
Items	11	9.4977	67.8569**	
Error	1441	.1400		
Total	1583			

** $p < .01$

Table 20

ANOVA and Reliability Table for Test II, Grade 7

Source of Variation	d.f.	MS	F	R
Individuals	131	.5362	3.7346**	0.73
Items	11	1.0058	69.9546**	
Error	1441	.1438		
Total	1583			

** $p < .01$

Table 21

Item Analysis for Test II for Grades Four, Five, Six and Seven

Item	Grade	Item Difficulty	r_b	X_{50}	β
13	4	.64	.94	- .39	2.70
	5	.69	.78	- .63	1.25
	6	.85	.73	-1.41	1.07
	7	.83	.80	-1.16	1.37
14	4	.33	.80	.54	1.31
	5	.47	.83	.09	1.51
	6	.62	.86	- .36	1.70
	7	.73	.78	- .80	1.24
15	4	.30	.81	.63	1.40
	5	.42	.75	.28	1.13
	6	.57	.82	- .21	1.45
	7	.65	.83	- .47	1.46
16	4	.54	.78	- .12	1.26
	5	.58	.73	- .29	1.06
	6	.64	.64	- .54	.84
	7	.76	.42	-1.66	.46
17	4	.33	.60	.71	.75
	5	.33	.67	.64	.91
	6	.54	.75	- .12	1.15
	7	.60	.82	- .30	1.44
18	4	.13	.97	1.16	4.18
	5	.16	.99	1.00	17.74
	6	.33	.88	.49	1.86
	7	.40	.88	.28	1.81
19	4	.21	.34	2.31	.37
	5	.21	.68	1.18	.92
	6	.30	.73	.70	1.07
	7	.33	.49	.92	.56
20	4	.01	.74	3.29	1.09
	5	.02	.63	3.44	.81
	6	.02	.40	5.36	.44
	7	.08	1.02	1.41	0

Table 21 (continued)

Item	Grade	Difficulty	r_b	X_{50}	β
21	4	.16	.32	3.07	.34
	5	.20	.52	1.58	.61
	6	.33	.58	.74	.72
	7	.35	.43	.90	.48
22	4	.20	.34	2.43	.36
	5	.16	.56	1.79	.67
	6	.22	.41	1.89	.45
	7	.24	.52	1.34	.61
23	4	.02	.76	2.80	1.03
	5	.04	.93	1.92	2.46
	6	.05	.84	2.01	1.55
	7	.10	.97	1.33	3.74
24	4	.02	.90	2.40	2.11
	5	.03	1.08	1.74	0
	6	.05	.67	2.41	.91
	7	.10	.97	1.37	4.20

Reliability Studies of Test III (Quantification of Probabilities)

Test III is a ten item test [67] on the concept of quantification of probabilities. Each item is a multiple choice item with three possible response options, only one of which is the correct response for the item. Each item was scored either right or wrong. The range of total scores possible for Test III is from 0 to 10 inclusive.

Tables 22 and 23 present the frequency distributions, mean scores and standard deviations for Test III. The scores were not corrected for guessing so the expected mean score, selecting options on a purely random basis, is 3.33.

The reliability coefficients for Test III, shown in Tables 24-27, are .67, .67, .70 and .80 for grades four, five,six and seven respectively. These reliabilities are modest but indicate that the items are reasonably consistent. Tables 24-27 also show that the differences among individuals and items are significant for all grades.

Table 28 gives the item analysis for Test III. Item 25 was an easy item for all pupils but it does not correlate very highly with the total test scores and is not a good discriminator. Items 26-34 were very difficult for fourth grade pupils. A difficulty index of .33 is what would be expected from a random selection of the response options. Items 26-34 functioned much the same for fifth grade pupils. Items 26 and 30-34 were very difficult with items 27-29 only slight easier. Item difficulties for items 26, 28, 29, 33 and 34 are similar for sixth and seventh grades. Item 27 was easiest for sixth grade while items

[67] See Appendix A for the test items. Test III consists of items 25-34.

30, 31 and 32 were easiest for seventh grade. Item 30 was an unusually difficult item for all grades. The errors made on this item, as well as errors on the other items, are discussed in Chapter V.

All items except item 1 have high item criterion correlations and appear to be good discriminators with β's of .61 or greater.

Table 22

Frequency Distributions of Total Scores on Test III for Grades Four, Five, Six and Seven

Total Score	Frequency Grade 4	Grade 5	Grade 6	Grade 7
10	2	4	3	11
9	1	6	14	16
8	4	7	12	28
7	5	13	20	9
6	15	11	16	10
5	20	28	22	17
4	11	20	14	12
3	24	19	14	6
2	21	15	11	13
1	23	6	5	9
0	6	3	1	1

Table 23

Mean Scores and Standard Deviations for Test III

Grade	Mean Score	Standard Deviation
4	3.58	2.26
5	4.70	2.31
6	5.48	2.38
7	5.96	2.79

Table 24

ANOVA and Reliability Table for Test III, Grade 4

Source of Variation	d.f.	MS	F	R
Individuals	131	.5143	3.0096**	0.67
Items	9	4.0380	23.6282**	
Error	1179	.1709		
Total	1319			

** $p < .01$

Table 25

ANOVA and Reliability Table for Test III, Grade 5

Source of Variation	d.f.	MS	F	R
Individuals	131	.5394	3.0340**	0.67
Items	9	5.4340	30.5671**	
Error	1179	.1778		
Total	1319			

** $p < .01$

Table 26

ANOVA and Reliability Table for Test III, Grade 6

Source of Variation	d.f.	MS	F	R
Individuals	131	.5702	3.3362**	0.70
Items	9	5.6449	33.0294**	
Error	1179	.1709		
Total	1319			
** $p < .01$				

Table 27

ANOVA and Reliability Table for Test III, Grade 7

Source of Variation	d.f.	MS	F	R
Individuals	131	.7854	4.9318**	0.80
Items	9	3.0169	18.9454**	
Error	1179	.1592		
Total	1319			
** $p < .01$				

Table 28

Item Analysis for Test III for Grades Four, Five, Six and Seven

Item	Grade	Item Difficulty	r_b	X_{50}	β
25	4	.77	.46	-1.61	.51
	5	.87	.29	-3.87	.31
	6	.89	.37	-3.31	.41
	7	.90	.34	-3.84	.36
26	4	.32	.73	.65	1.06
	5	.41	.74	.29	1.09
	6	.51	.76	- .02	1.16
	7	.58	.95	- .20	3.21
27	4	.38	.52	.58	.61
	5	.58	.57	- .35	.69
	6	.79	.69	- .74	.96
	7	.62	.93	- .33	2.48
28	4	.48	.66	.07	.88
	5	.63	.66	- .51	.87
	6	.73	.77	- .81	1.22
	7	.70	.86	- .62	1.68
29	4	.34	.71	.58	1.01
	5	.56	.54	- .29	.64
	6	.67	.67	- .63	.91
	7	.67	.79	- .55	1.27
30	4	.13	.55	2.05	.66
	5	.15	.69	1.49	.97
	6	.16	.75	1.33	1.12
	7	.32	.76	.62	1.18
31	4	.25	.82	.81	1.45
	5	.29	.84	.67	1.57
	6	.42	.68	.31	.92
	7	.58	.75	- .28	1.13
32	4	.30	.77	.70	1.19
	5	.35	.70	.56	.99
	6	.41	.76	.28	1.16
	7	.56	.71	- .21	1.00

Table 28 (continued)

Item	Grade	Difficulty	r_b	X_{50}	β
33	4	.28	.87	.66	1.77
	5	.41	.74	.32	1.10
	6	.51	.77	-.02	1.23
	7	.55	.87	-.15	1.76
34	4	.34	.55	.75	.66
	5	.45	.72	.17	1.05
	6	.48	.54	.10	.65
	7	.48	.59	.10	.73

Reliability Studies of Subtest I-A (Sample Space : Simple Counting)

Subtest I-A is a six item test consisting of the first six items of Test I. These six items involved only simple counting to determine all of the outcomes of an experiment. Tables 29 and 30 give the frequency distributions, mean scores and standard deviations for this subtest. The high mean scores reported in Table 30 indicate that this subtest was relatively easy for all grades.

The analysis of variance and Hoyt reliability coefficients for Subtest I-A are reported in Tables 31-34. The ANOVA tables show that the differences between individuals and the differences between items are significant for each grade level. The reliability coefficients for grades four, five, six and seven are .58, .60, .63 and .66 respectively. These are good reliabilities for a six item test.

The item analysis for this subtest is presented in Table 35. The total score for the first six items was used as the criterion variable for this analysis. The item difficulties are the same as those reported in Table 14 for these items. The item criterion correlations are generally very high. All items appear to be very good discriminators with all but two of the β's greater than .87. The fact that these items were easy items for all grades is reflected by the X_{50}'s, all of which are either below or very near the mean of the criterion scores.

Table 29

Frequency Distributions of Total Scores on Subtest I-A for Grades Four, Five, Six and Seven

Total Score	Frequency Grade 4	Grade 5	Grade 6	Grade 7
6	28	38	61	67
5	19	17	19	30
4	19	41	28	17
3	24	17	11	9
2	23	11	11	6
1	13	5	2	1
0	6	3	0	2

Table 30

Mean Scores and Standard Deviations for Subtest I-A

Grade	Mean Score	Standard Deviation
4	3.55	1.79
5	4.20	1.55
6	4.77	1.40
7	5.00	1.35

Table 31

ANOVA and Reliability Table for Subtest I-A, Grade 4

Source of Variation	d.f.	MS	F	R
Individuals	131	.5391	3.1572**	0.68
Items	5	1.8379	10.7643**	
Error	655	.1707		
Total	791			

** $p < .01$

Table 32

ANOVA and Reliability Table for Subtest I-A, Grade 5

Source of Variation	d.f.	MS	F	R
Individuals	131	.4027	2.4720**	0.60
Items	5	1.4316	8.7885**	
Error	655	.1629		
Total	791			

** $p < .01$

Table 33

ANOVA and Reliability Table for Subtest I-A, Grade 6

Source of Variation	d.f.	MS	F	R
Individuals	131	.3272	2.6694**	0.63
Items	5	1.1424	9.3201**	
Error	655	.1226		
Total	791			

** $p < .01$

Table 34

ANOVA and Reliability Table for Subtest I-A, Grade 7

Source of Variation	d.f.	MS	F	R
Individuals	131	.3053	2.9543**	0.66
Items	5	.4606	4.4566**	
Error	655	.1034		
Total	791			

** $p < .01$

Table 35

Item Analysis for Subtest I-A for Grades Four, Five, Six and Seven

Item	Grade	Item Difficulty	r_b	X_{50}	β
1	4	.67	.66	- .66	.87
	5	.81	.75	-1.17	1.14
	6	.94	.51	-3.07	.59
	7	.92	.86	-1.65	1.73
2	4	.71	.82	- .68	1.46
	5	.71	.76	- .74	1.18
	6	.72	.91	- .64	2.20
	7	.80	.91	- .91	2.18
3	4	.68	.74	- .62	1.09
	5	.71	.71	- .76	1.02
	6	.73	.91	- .69	2.20
	7	.79	.97	- .82	3.79
4	4	.58	.65	- .29	.86
	5	.79	.75	-1.07	1.12
	6	.86	.64	-1.71	.84
	7	.89	.73	-1.66	1.06
5	4	.50	.94	- .01	2.99
	5	.65	.77	- .48	1.22
	6	.81	.97	- .91	3.75
	7	.83	1.04	- .89	0
6	4	.42	.95	.22	3.21
	5	.53	.85	- .08	1.58
	6	.70	.94	- .57	2.86
	7	.78	.93	- .83	2.50

Reliability Studies of Subtest I-B (Sample Space : Combinations)

Subtest I-B is a six item test consisting of the last six items in Test I. All of these items involved combinations. Tables 36 and 37 present the frequency distributions, mean scores and standard deviations for Subtest I-B. Comparing these tables with Tables 29 and 30 points out that Subtest I-B was more difficult than Subtest I-A. This result is not surprising since it was expected that combinations would be more difficult than simple counting for elementary school children.

Although Subtest I-B was more difficult than Subtest I-A the reliabilities for Subtest I-B are greater than the reliabilities for Subtest I-A. The reliabilities for Subtest I-B are given in Tables 38-41 along with the ANOVA tables. The reliabilities for this subtest are .81, .81, .74 and .78 for grades four, five, six and seven respectively. These are very good reliabilities for a six item test.

The item analysis for Subtest I-B is given in Table 42. Items 7-10 have high item-criterion correlations and are good discriminators. Little can be said about items 11 and 12 because the r_b values are greater than 1 and corresponding β's are zero. It is easy to see from Table 36 that the total scores on this subtest are not normally distributed. As reported in a previous section of this chapter, r_b values greater than unity may result when the assumption of normality is violated. When an r_b greater than 1 is computed for an item the value of β for that item is set equal to zero.

Table 36

Frequency Distributions of Total Scores on Subtest I-B for Grades Four, Five, Six and Seven

Total Score	Frequency Grade 4	Grade 5	Grade 6	Grade 7
6	10	26	28	34
5	10	12	37	36
4	13	16	13	18
3	9	22	13	16
2	21	14	24	10
1	25	19	9	4
0	44	23	8	14

Table 37

Mean Scores and Standard Deviations for Subtest I-B

Grade	Mean Score	Standard Deviation
4	1.92	1.95
5	2.96	2.12
6	3.80	1.87
7	4.00	1.92

Table 38

ANOVA and Reliability Table for Subtest I-B, Grade 4

Source of Variation	d.f.	MS	F	R
Individuals	131	.6403	5.3163**	0.81
Items	5	2.0137	16.7203**	
Error	655	.1204		
Total	791			

** $p < .01$

Table 39

ANOVA and Reliability Table for Subtest I-B, Grade 5

Source of Variation	d.f.	MS	F	R
Individuals	131	.7561	5.2975**	0.81
Items	5	1.0937	7.6633**	
Error	655	.1427		
Total	791			

** $p < .01$

Table 40

ANOVA and Reliability Table for Subtest I-B, Grade 5

Source of Variation	d.f.	MS	F	R
Individuals	131	.5871	3.8467**	0.74
Items	5	1.4386	9.4256**	
Error	655	.1526		
Total	791			

** $p < .01$

Table 41

ANOVA and Reliability Table for Subtest I-B, Grade 7

Source of Variation	d.f.	MS	F	R
Individuals	131	.6209	4.6093**	0.78
Items	5	1.2879	9.5612**	
Error	655	.1347		
Total	791			

** $p < .01$

Table 42

Item Analysis for Subtest I-B for Grades Four, Five, Six and Seven

Item	Grade	Item Difficulty	r_b	X_{50}	β
7	4	.56	.88	- .17	1.85
	5	.67	.76	- .57	1.19
	6	.81	.79	-1.11	1.29
	7	.77	.86	- .87	1.69
8	4	.31	.93	.53	2.60
	5	.49	.87	.03	1.73
	6	.67	.73	- .61	1.08
	7	.68	.74	- .64	1.10
9	4	.26	.93	.69	2.53
	5	.41	.88	.27	1.83
	6	.51	.57	- .03	.67
	7	.49	.77	.02	1.21
10	4	.29	.95	.59	3.07
	5	.45	.97	.13	3.90
	6	.56	.98	- .15	4.61
	7	.62	.95	- .32	3.12
11	4	.24	1.03	.67	0
	5	.47	.97	.07	4.26
	6	.62	1.02	- .30	0
	7	.72	1.10	- .53	0
12	4	.26	1.01	.64	0
	5	.47	.97	.07	4.26
	6	.62	1.03	- .30	0
	7	.71	1.05	- .53	0

Reliability Studies of Subtest II-A (Probability: Simple Counting)

Subtest II-A is a six item test consisting of the first six items of Test II. These items involve the probability of a simple event which is a subset of a sample space obtained by simple counting. The frequency distributions, mean scores and standard deviations for this subtest are given in Tables 43 and 44. The distributions for grades four and five are quite similar with the mean score for grade five only slightly higher than the mean score for grade four. The distributions for grades six and seven are also somewhat similar to each other.

Tables 45-48 present the ANOVA tables and reliability coefficients for Subtest II-A. The reliabilities are almost identical for all grades. The reliabilities are .74, .74, .73 and .75 for grades four through seven respectively. The ANOVA tables show that individual and item differences are highly significant for each grade level.

The item analysis for Subtest II-A is reported in Table 49. In general these items correlate well with the criterion scores and are good discriminators for all grades.

Table 43

Frequency Distributions of Total Scores on Subtest II-A for Grades Four, Five, Six and Seven

Total Score	Frequency			
	Grade 4	Grade 5	Grade 6	Grade 7
6	10	13	27	38
5	9	12	22	23
4	14	20	18	21
3	17	17	19	17
2	30	29	27	18
1	29	23	13	10
0	23	18	6	5

Table 44

Mean Scores and Standard Deviations for Subtest II-A

Grade	Mean Score	Standard Deviation
4	2.28	1.79
5	2.65	1.85
6	3.55	1.83
7	3.97	1.81

Table 45

ANOVA and Reliability Table for Subtest II-A, Grade 4

Source of Variation	d.f.	MS	F	R
Individuals	131	.5402	3.7781**	0.74
Items	5	4.4346	31.0127**	
Error	655	.1430		
Total	791			
** p < .01				

Table 46

ANOVA and Reliability Table for Subtest II-A, Grade 5

Source of Variation	d.f.	MS	F	R
Individuals	131	.5750	3.8839**	0.74
Items	5	4.6051	31.1041**	
Error	655	.1481		
Total	791			
** p < .01				

Table 47

ANOVA and Reliability Table for Subtest II-A, Grade 6

Source of Variation	d.f.	MS	F	R
Individuals	131	.5633	3.7147**	0.73
Items	5	3.6697	24.2015**	
Error	655	.1516		
Total	791			

** $p < .01$

Table 48

ANOVA and Reliability Table for Subtest II-A, Grade 7

Source of Variation	d.f.	MS	F	R
Individuals	131	.5520	4.0151**	0.75
Items	5	2.9899	21.7476**	
Error	655	.1375		
Total	791			

** $p < .01$

Table 49

Item Analysis for Subtest II-A for Grades Four, Five, Six and Seven

Item	Grade	Item Difficulty	r_b	X_{50}	β
13	4	.64	.91	- .40	2.16
	5	.69	.85	- .58	1.64
	6	.85	.77	-1.33	1.22
	7	.83	.85	-1.10	1.60
14	4	.33	.95	.45	2.97
	5	.47	.97	.07	4.27
	6	.62	.96	- .32	3.40
	7	.73	.94	- .66	2.79
15	4	.30	.98	.53	4.44
	5	.42	.88	.24	1.88
	6	.57	.93	- .18	2.47
	7	.65	1.00	- .39	0
16	4	.54	.71	- .13	1.02
	5	.58	.81	- .26	1.36
	6	.64	.68	- .51	.94
	7	.76	.61	-1.13	.78
17	4	.33	.72	.60	1.02
	5	.33	.71	.60	1.01
	6	.54	.80	- .12	1.36
	7	.60	.94	- .26	2.86
18	4	.13	1.07	1.06	0
	5	.16	.99	1.00	38.05
	6	.33	.94	.45	2.84
	7	.40	.92	.27	2.43

Reliability Studies of Subtest II-B (Probability: Combinations)

Subtest II-B is a six item test consisting of the last six items of Test II. These items involve the probability of a simple event which is a subset of a sample space obtained by using ideas of combinations. Tables 50 and 51 present the frequency distributions, mean scores and standard deviations for this subtest. An examination of these tables indicates that this subtest was extremely difficult for all subjects. Only three children answered all of the items correctly. More than 50% of the children in grades four and five and more than 30% of the children in grades six and seven answered all of the items incorrectly.

The ANOVA tables and reliabilities for this subtest are reported in Tables 52-55. Since the items were extremely difficult the reliabilities are very low. The reliabilities for grades four, five, six and seven are .18, .58, .41 and .62 respectively. In grade four the differences among individuals is not significant while the difference among items is significant. In the other grades both the differences among individuals and among items are significant.

The item analysis for Subtest II-B is given in Table 56. As previously noted all items were very difficult. Since the distributions of total scores are obviously not normal distributions, as can be seen from Table 50, it is not surprising that many of the r_b values for this subtest are greater than unity. All of the X_{50}'s are above the mean of the criterion score because of the difficulty of the items. Items 19, 21 and 22 appear to be good discriminators among pupils who do well on this subtest. Little can be said about the items 20, 23 and 24 because many of the β's were set equal to zero. A more detailed discussion of the types of errors children made on these items is presented in Chapter V.

Table 50

Frequency Distributions of Total Scores on Subtest II-B for Grades Four, Five, Six and Seven

Total Score	Frequency Grade 4	Grade 5	Grade 6	Grade 7
6	0	1	0	2
5	0	1	2	4
4	0	1	1	3
3	1	3	6	7
2	21	15	26	24
1	37	33	44	43
0	75	78	53	49

Table 51

Mean Scores and Standard Deviations for Subtest II-B

Grade	Mean Score	Standard Deviation
4	.62	.77
5	.66	1.03
6	.97	1.04
7	1.18	1.34

Table 52

ANOVA and Reliability Table for Subtest II-B, Grade 4

Source of Variation	d.f.	MS	F	R
Individuals	131	.1006	1.2222	0.18
Items	5	1.2838	15.5973**	
Error	655	.0823		
Total	791			

** $p < .01$

Table 53

ANOVA and Reliability Table for Subtest II-B, Grade 5

Source of Variation	d.f.	MS	F	R
Individuals	131	.1777	2.3957**	0.58
Items	5	1.1174	15.0663**	
Error	655	.0742		
Total	791			

** $p < .01$

Table 54

ANOVA and Reliability Table for Subtest II-B, Grade 6

Source of Variation	d.f.	MS	F	R
Individuals	131	.1831	1.7083**	0.41
Items	5	2.6293	24.5372**	
Error	655	.1072		
Total	791			

** $p < .01$

Table 55

ANOVA and Reliability Table for Subtest II-B, Grade 7

Source of Variation	d.f.	MS	F	R
Individuals	131	.3023	2.6240**	0.62
Items	5	2.0394	17.6999**	
Error	655	.1152		
Total	791			

** $p < .01$

Table 56

Item Analysis for Subtest II-B for Grades Four, Five, Six and Seven

Item	Grade	Item Difficulty	r_b	X_{50}	β
19	4	.21	.93	.87	2.53
	5	.21	1.05	.76	0
	6	.30	.86	.60	1.65
	7	.33	.80	.56	1.34
20	4	.01	1.12	2.17	0
	5	.02	1.10	1.97	0
	6	.02	.77	2.80	1.22
	7	.08	1.23	1.16	0
21	4	.16	.92	1.08	2.43
	5	.20	.99	.83	6.86
	6	.33	.80	.53	1.35
	7	.35	.67	.58	.89
22	4	.20	.60	1.39	.74
	5	.16	.82	1.21	1.45
	6	.22	.61	1.27	.77
	7	.24	.71	.99	1.00
23	4	.02	.93	2.14	2.60
	5	.04	1.22	1.45	0
	6	.05	1.23	1.38	0
	7	.10	1.25	1.03	0
24	4	.02	.98	2.25	3.68
	5	.03	1.43	1.31	0
	6	.05	1.15	1.40	0
	7	.10	1.17	1.15	0

Summary

Table 57 gives the reliability coefficients for each test and subtest for each grade. The reliabilities for Tests I, II and III are very good considering each test had twelve or fewer items. These reliabilities indicate that the items on each test are fairly consistent and suggest that the results of these tests may be interpreted with a high degree of confidence. The reliabilities for Subtests I-A, I-B and II-A are also very good and give support to the interpretation of results contained in Chapter V. The extreme difficulty of the items on Subtest II-B resulted in a very low reliability for grade four and a poor reliability for grade six.

In general all items on Tests I, II and III are very good discriminators at the X_{50} points for each grade. Only 8% of the β's reported for all grades for all items are less than .60. A β of .60 represents a slope of approximately 31 degrees for the item characteristic curve at the X_{50} point.

Using the total score of each test as criterion, 75% of the X_{50}'s on Test I are below the mean, 27% of the X_{50}'s on Test II are below the mean and 48% of the X_{50}'s on Test III are below the mean. Thus Test II was the most difficult test, with most of the difficulty being accounted for by the last six items, or Subtest II-B.

Table 57

Summary Table of Reliability Coefficients of the Probability Concept Tests and Subtests for Grades Four, Five, Six and Seven

Test	Number Of Items	Grade	Reliability
I Sample Space	12	4	.81
		5	.81
		6	.76
		7	.81
II Probability of a Simple Event	12	4	.62
		5	.72
		6	.72
		7	.73
III Quantification of Probability	10	4	.67
		5	.67
		6	.70
		7	.80
I-A Sample Space : Simple Counting	6	4	.68
		5	.60
		6	.63
		7	.66
I-B Sample Space : Combinations	6	4	.81
		5	.81
		6	.74
		7	.78
II-A Probability : Simple Counting	6	4	.74
		5	.74
		6	.73
		7	.75
II-B Probability : Combinations	6	4	.18
		5	.58
		6	.41
		7	.62

Chapter V

ANALYSIS OF THE DATA

The results of the analysis of data for this study are presented in three parts. The first part deals with the testing of hypotheses regarding the relationship between the combined performance scores on the three probability tests and the factors of I.Q., sex and grade. The second part is concerned with the relationship between performances on each of the three tests, and four subtests, within each grade. The third part presents an analysis of the incorrect responses made on each of the test items.

Part 1: The Testing of Hypotheses

In the Statement of the Problem, Chapter II, seven hypotheses were presented for testing. To test these hypotheses a multivariate analysis of covariance was run to test for equality of mean vectors over the factors under consideration. The dependent variables were the performance scores on the three tests of probability concepts. The covariates were the grade equivalent scores on the three parts of the Stanford Arithmetic Achievement Tests: computation, concepts and applications. Since the subjects for this study were randomly selected and the results of a standardized test used as covariants it seemed reasonable to presume that the assumptions underlying the analysis of covariance had been met.

The application of analysis of covariance may be used to increase the precision in a randomized experiment. Cochran indicates that gain in precision is directly related to the size of the correlation coefficient p between the covariant and dependent variable. He goes on to say, "If p is less than 0.3 in absolute value, the reduction in variance is inconsequential, but as p mounts toward unity, sizable increases in precision are obtained."[68] The multiple correlations between the grade equivalent scores on the Stanford Arithmetic Achievement tests (covariates) and the raw scores on the three probability tests (dependent variables) are .38, .40 and .30. (See Table 59). Since these correlations are all greater than or equal to .3 it was decided to include the arithmetic achievement scores as covariates in the study.

The data analyses for this part were done using a computer program written by J. Finn[69] and adapted for use on the CDC 3600 computer at the University of Wisconsin. This program used the computational procedures outlined by Bock.[70]

For each hypothesis the computer program calculated an overall F ratio for the multivariate test of equality of mean vectors as well as

---

[68] William G. Cochran, "Analysis of Covariance: Its Nature and Uses," Biometrics Vol. 13, No. 3 (September, 1957), pp. 262-263.

[69] Jeremy D. Finn, Multivariance: Fortran Program for Univariate and Multivariate Analysis of Variance and Covariance (Buffalo: State University of New York at Buffalo, Department of Educational Psychology, May, 1967).

[70] R. Darrell Bock, "Programming Univariate and Multivariate Analysis of Variance," Technometrics, 5 (February, 1963), pp. 95-117.

univariate F statistics for the dependent variables. A discriminant function analysis for each between-cell hypothesis was also computed. This discriminant function is reported and discussed for each hypothesis test in which the overall F ratio indicated that the variation among mean vectors was significant ($p < .01$). In addition to calculating the F statistics for each hypothesis the program also performed a multivariate regression analysis to test the covariates' relationship with the dependent variables.

Table 58 presents the multivariate regression statistics which summarize the contributions of the covariants to the analysis. In a multivariate sense, the three covariants have a significant association with the dependent variables. The multivariate test provided a chi square value of 143.98 with 9 degrees of freedom which is significant at the .01% level. The univariate statistics show that the multiple correlations of the three covariates with Tests I, II and III are .38, .40 and .30 respectively. The corresponding F ratios indicate that these correlations differ significantly from zero. This indicates that each of the dependent variables can have a small but significant amount of its variance predicted by the three covariates. The square of the multiple correlation coefficient is an approximation of the per cent of variance predicted by the covariates. The approximate amounts of variance that can be predicted are 15%, 16% and 9% for Tests I, II and III respectively.

Table 58

Statistics for Regression Analysis with Three Covariates

Variable	Square Multiple Correlation	Multiple Correlation	F	P Less Than
Test I	.15	.38	28.68	.0001
Test II	.16	.40	31.95	.0001
Test III	.09	.30	16.52	.0001

Degrees of Freedom for Hypothesis = 3
Degrees of Freedom for Error = 501

Chi Square Test of Hypothesis of No Association Between Dependent and Independent Variables = 143.98

Degrees of Freedom = 9 $p < .0001$

The results of the statistical testing of the hypotheses are presented and discussed individually.

Hypothesis 1: There is no difference in the mean performances of children in the three I.Q. groups.

The mean scores on the three probability tests for the three I.Q. groups are displayed in Table 59. The multivariate and univariate F statistics for the main effect of I.Q. are given in Table 60.

An examination of Table 60 shows that the first hypothesis can be rejected. The multivariate F statistic indicates that there is significant variation among the mean vectors for the I.Q. groups. The univariate F's indicate that the mean differences for each of the three probability tests are highly significant. Therefore, the variation among mean vectors is due to the significant differences among each of the three elements (dependent variables) of the vectors.

Since the variation among mean vectors for the main effect of I.Q. was significant one discriminant function for this effect is also reported in Table 60. The multivariate results for the main effect of I.Q. indicated that only one dimension of discrimination was significant. The discriminant function is the linear function of the dependent variables which maximally discriminates among groups in a least squares sense. The discriminant function provides a means of characterizing the multivariant differences between groups.

The discriminant function for I.Q. groups is:

$$V_{IQ} = .6162T_1 + .3532T_2 + .4572T_3$$

The coefficients are presented in standardized form and thus represent the relative magnitude of the contribution of each dependent variable to the discrimination between groups. Since the weights in the function are all positive the discrimination between I.Q. groups is an overall effect. An inspection of Table 59 shows that the direction of mean differences among I.Q. groups is the same for each of the dependent variables. The highest I.Q. group has the highest performance scores and the lowest I.Q. group the lowest performance scores. The largest contribution to the discriminant function is Test I and Test II contributes the least.

The distributions of standardized discriminant scores, V_{IQ}, for the 176 pupils in each I.Q. groups appear in Figure 1. This figure clearly illustrates that the best discrimination is between the low I.Q. group and the high I.Q. group. The distribution of the low group is substantially lower than the distribution for the high group. The distribution of the middle group overlaps a considerable portion of each of the other distributions. This indicates that the function does not discriminate very well between the low-middle groups and the middle-high groups.

These observations are consistent with Leake's results in which he concluded that mental age was a significant factor in the performances of junior high students on three probability tests. They are also consistent with the findings of Pire.

Table 59

Mean Scores of the I.Q. Groups

I.Q.	Raw Score Means Test I	Test II	Test III	Adjusted Means Test I	Test II	Test III
Range I (72 -104)	5.56	2.67	3.59	2.39	-.39	1.22
Range II (105-113)	7.69	3.88	4.92	2.99	-.37	1.49
Range III (114-144)	9.44	5.35	6.25	3.06	-.28	1.80

Table 60

MANOVA Table for Hypothesis 1

F-Ratio for Multivariate Test of Equality of Mean Vectors = 11.9445

Degrees of Freedom for Hypothesis = 6

Degrees of Freedom for Error = 998.00

$p < .001$

Variable	Between Mean Square	Univariate F	P Less Than
Test I	129.83	22.85	.0001
Test II	48.58	13.41	.0001
Test III	68.77	15.80	.0001

Degrees of Freedom for Hypothesis = 2

Degrees of Freedom for Error = 501

3 Covariates Have Been Eliminated

Discriminant Function, I.Q.:

$$V_{IQ} = + .6162T_1 + .3532T_2 + .4572T_2$$

Hypothesis 2: There is no difference in the mean performances of boys and girls.

Table 61 presents the mean scores for boys and girls on each of the three tests. In Table 62 appears the multivariate and univariate F statistics for the main effect of sex.

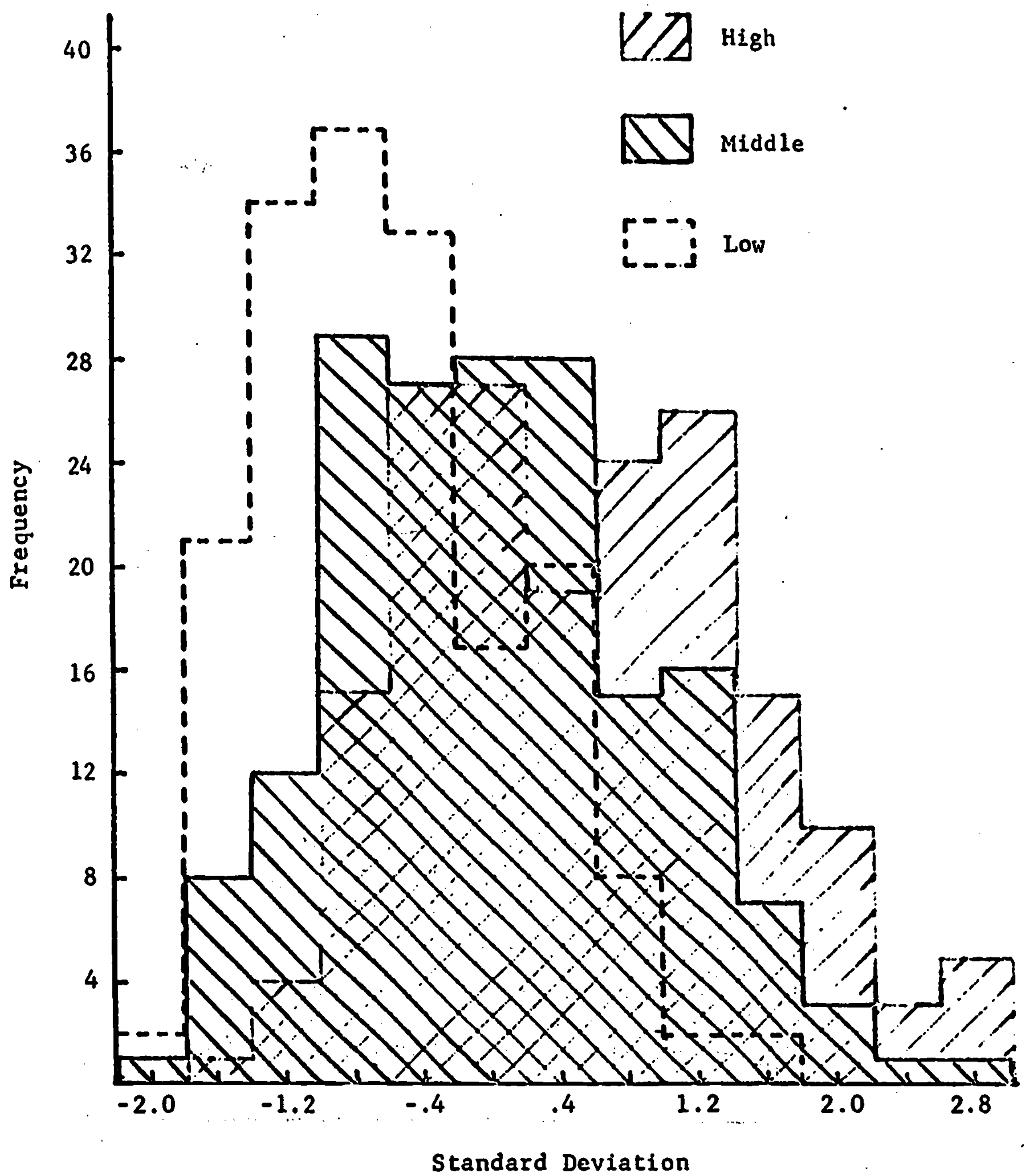


Figure 1. Distributions of Discriminant Scores for I.Q. Groups

The multivariate F statistic reported in Table 62 indicates that the variation among mean vectors for boys and girls is significant. Therefore, in a multivariate sense, hypothesis 2 can be rejected. An examination of the univariate results reported in Table 62 helps explain the source of this variation. The probability levels associated with the mean differences on Tests I and III are .0164 and .0382 respectively. The F ratio for Test II is less than 1 indicating that the difference between the mean scores on this test is not significant. Therefore the significant variation among mean vectors can be attributed mainly to the variations among the two elements of the vectors corresponding to the mean scores on Tests I and III.

An inspection of Table 61 shows that when the mean scores are adjusted for the covariates the direction of the mean difference on Test III is reversed. Since this difference is very small and the mean differences of the covariates are also very small (less than .2) this change in direction may be due to chance. The direction of the adjusted mean differences on Tests I and III show that the performances of the girls was slightly better than the performances of the boys on these tests. On Test II the adjusted mean for boys is higher than the adjusted mean for girls but their difference is clearly not significant.

Since there is only one degree of freedom for the main effect of sex there is only one characteristic root associated with this factor. Therefore there can only be one dimension of discrimination for the main effect of sex. The discriminant function for this factor, as reported in Table 62, is:

$$V_{sex} = .8485T_1 - .1374T_2 - .7199T_3.$$

The standardized coefficients indicate that the discrimination between boys and girls is accounted for primarily by Tests I and III. Since the first coefficient is positive and the other two are negative the discrimination is a contrast of sample space with probability of an event.

Figure 2 displays the distributions of standardized discriminant scores, V_{sex}, for 264 boys and 264 girls included in the sample. The distributions have a considerable overlap with the distribution for girls being slightly higher than the distribution for boys. The distributions displayed in Figure 2 indicate that the function does not discriminate very well between boys and girls.

Table 61

Mean Scores of Boys and Girls

Sex	Raw Score Means			Adjusted Means		
	Test I	Test II	Test III	Test I	Test II	Test III
Boys	7.35	3.99	5.10	2.63	-.27	1.36
Girls	7.78	3.94	4.74	2.98	-.51	1.64

Table 62

MANOVA Table for Hypothesis 2

F-Ratio for Multivariate Test of Equality of Mean Vectors = 4.2819

Degrees of Freedom for Hypothesis = 3

Degrees of Freedom for Error = 499.00

$p < .0054$

Variable	Between Mean Square	Univariate F	P Less Than
Test I	33.01	5.81	.0164
Test II	0.48	0.13	.7169
Test III	18.81	4.32	.0382

Degrees of Freedom for Hypothesis = 1

Degrees of Freedom for Error = 501

3 Covariants Have Been Eliminated

Discriminant Function, Sex:

$$V_{sex} = .8485T_1 - .1374T_2 - .7199T_3$$

Hypothesis 3: There is no difference in the mean performances of children in the four grades.

The mean scores on the three probability tests for the four grades are given in Table 63. The multivariant and univariant F statistics for the main effect of grade appear in Table 64.

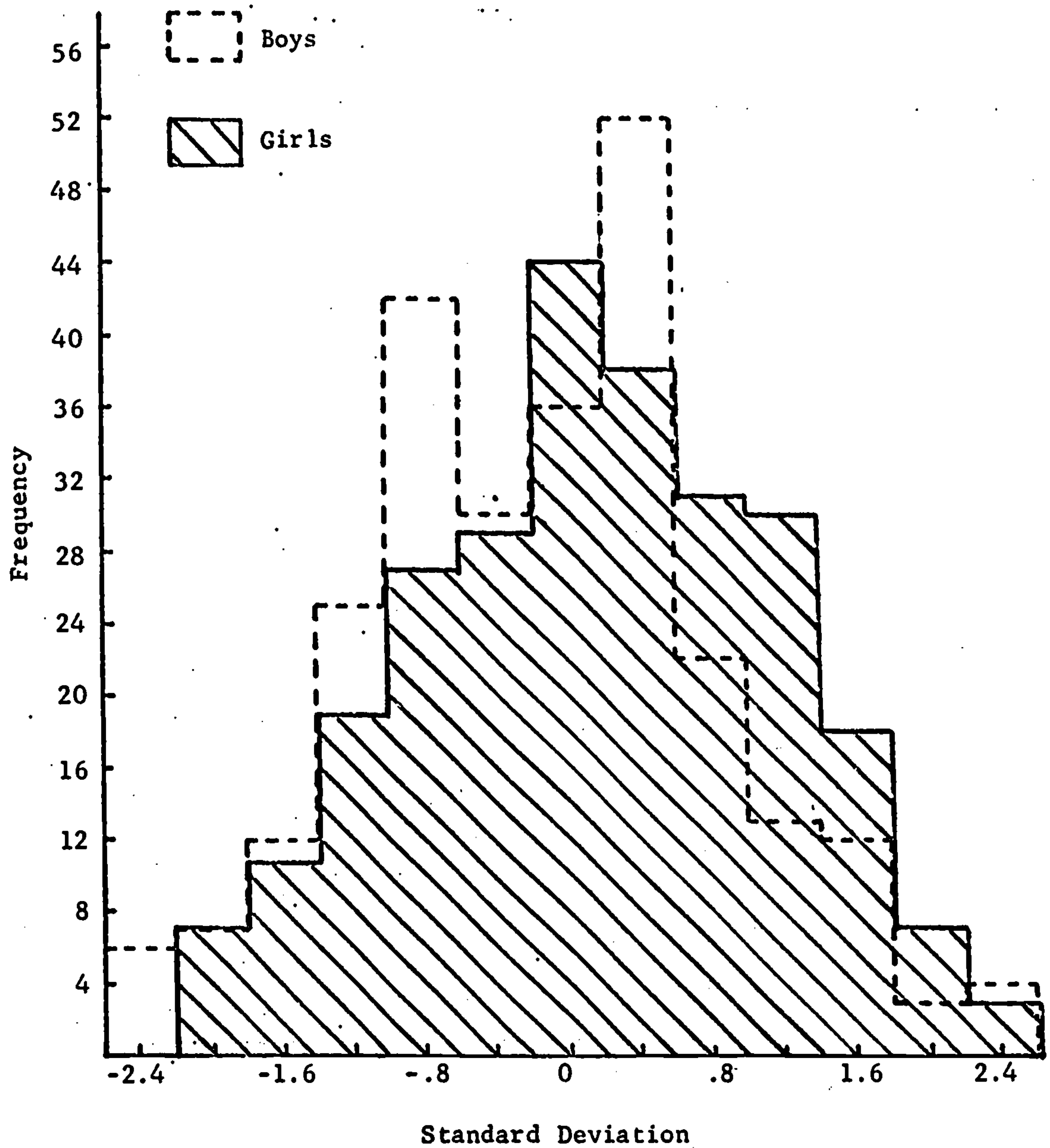


Figure 2. Distributions of Discriminant Scores for Boys and Girls.

The multivariate F statistic reported in Table 64 indicates that there is significant variation among the mean vectors for the grades. Therefore hypothesis 3 can be rejected. An examination of the univariate F statistics given in Table 64 indicates the mean differences for grades on Test I are significant ($p < .01$). The mean differences on Tests II and III are not significant. Therefore the significant variation among mean vectors for grades is due primarily to the significant differences among the elements of the vectors corresponding to the mean scores on Test I.

The multivariate statistics for hypothesis 3 showed that only one dimension of discrimination was significant. Therefore only one discriminant function for the main effect of grade is reported in Table 64. This function is:

$$V_{grade} = .8389T_1 - .4374T_2 + .4568T_3$$

The standardized coefficients indicate that Test I contributes most to the discrimination with Tests II and III having weights which are opposite in sign but of approximately the same magnitude. The discrimination between grades is accounted for by a contrast of Test I and Test III with Test II. This may be interpreted as a contrast between sample space and probability of an event.

Figure 3 presents a histogram showing the distributions of scores, V_{grade}, for the subjects in the sample. The sample included 132 children from each grade. The distributions for grades six and seven are very much alike. The distribution of grade four is substantially lower than the distributions for grades six and seven. The distribution for grade five overlaps a considerable portion of each of the other distributions. The function may be interpreted as showing the overall performance of grades six and seven is considerably higher than the overall performance of grade four and to a lesser degree better than the overall performance of grade five on these tests.

These observations are consistent with the findings of Leake.

Table 63

Mean Scores of the Four Grades

Grade	Raw Score Means			Adjusted Means		
	Test I	Test II	Test III	Test I	Test II	Test III
4	5.50	2.89	3.54	1.51	-1.03	.71
5	7.18	3.31	4.70	2.48	-.70	1.45
6	8.57	4.52	5.48	3.24	.15	1.57
7	9.00	5.15	5.96	4.01	.20	2.27

Table 64

MANOVA Table for Hypothesis 3

F-Ratio for Multivariate Test of Equality of Mean Vectors = 2.4627

Degrees of Freedom for Hypothesis = 9

Degrees of Freedom for Error = 1214.56

$p < .0088$

Variable	Between Mean Square	Univariate F	P Less Than
Test I	25.05	4.41	.0045
Test II	5.23	1.44	.2290
Test III	7.50	1.72	.1615

Degrees of Freedom for Hypothesis = 3

Degrees of Freedom for Error = 501

3 Covariants Have Been Eliminated

Discriminant Function, Grade

$$V_{grade} = .8389T_1 - .4374T_2 + .4568T_3$$

Hypothesis 4: There is no difference in the mean performances of children in the three I.Q. groups across the two sex groups.

Table 66 presents the multivariate and univariate F statistics for the interaction I.Q. x sex. The multivariate F ratio indicates that there is no significant interaction in the multivariate sense. Therefore hypothesis 4 can not be rejected. The univariate F ratios

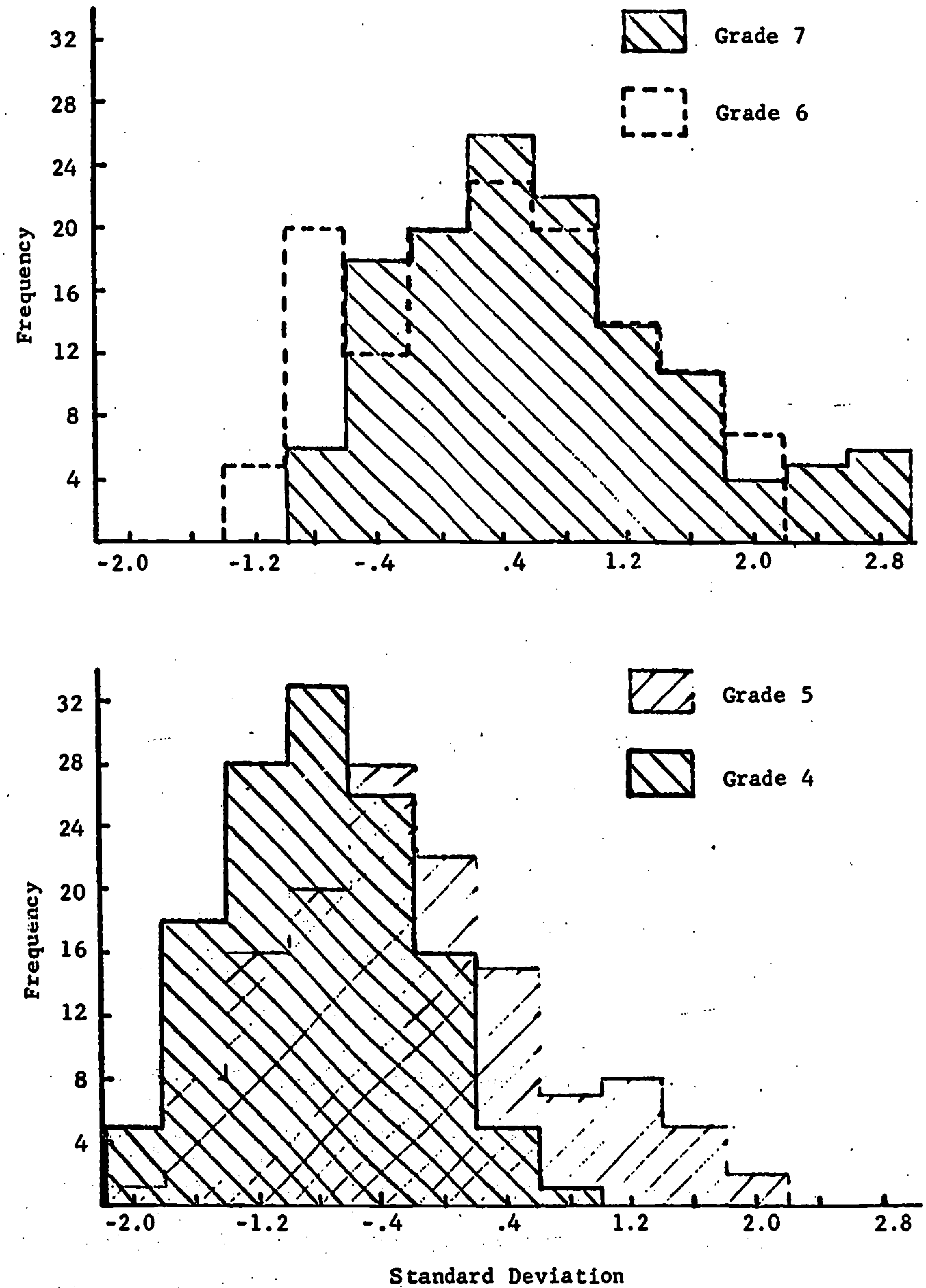


Figure 3. Distributions of Discriminant Scores for Grades

show the interaction of I.Q. x sex is not significant for Tests I and III but the probability level for the interaction on Test II is .017. An examination of the adjusted mean scores for Test II, presented in Table 65, shows that the girls in the low I.Q. group did better than the boys, but in the middle and high I.Q. groups the boys did better than the girls. On Tests I and III the adjusted mean scores for girls are higher than the adjusted mean scores for the boys in all three I.Q. groups.

Hypothesis 5: There is no difference in the mean performances of children in the three I.Q. groups across the four grade levels.

The mean scores for the interaction I.Q. x grade are displayed in Table 67. The multivariate and univariate F statistics for this interaction are presented in Table 68. In the multivariate sense there is no significant interaction of I.Q. x grade. Therefore hypothesis 5 can not be rejected. The univariate F's are all small indicating that there is no significant interaction on Tests I and II but the interaction on Test III approaches significance with a probability level of .03.

Table 65

Mean Scores of Interaction, I.Q. x Sex

IQ	Sex	Raw Score Means Test I	Test II	Test III	Adjusted Means Test I	Test II	Test III
1	M	5.40	2.68	3.68	2.25	-.41	1.19
	F	5.73	2.66	3.50	2.52	-.37	1.25
2	M	7.43	3.59	5.21	2.91	-.17	1.36
	F	7.94	4.16	4.63	3.07	-.56	1.62
3	M	9.22	5.69	6.42	2.74	-.23	1.53
	F	9.66	5.01	6.08	3.36	-.33	2.06

Table 66

MANOVA Table for Hypothesis 4

F-Ratio for Multivariate Test of Equality of Mean Vectors = 1.8510

Degrees of Freedom for Hypothesis = 6

Degrees of Freedom for Error = 998.00

$p < .0864$

Variable	Between Mean Square	Univariate F	P Less Than
Test I	1.66	.29	.7463
Test II	14.89	4.11	.0170
Test III	1.43	.33	.7195

Degrees of Freedom for Hypothesis = 2

Degrees of Freedom for Error = 501

3 Covariates Have Been Eliminated

Table 67

Mean Scores of Interaction, I.Q. x Grade

IQ	Grade	Raw Score Means Test I	Test II	Test III	Adjusted Means Test I	Test II	Test III
1	4	3.48	1.59	2.66	.87	-.48	.81
	5	5.09	1.93	3.21	2.24	-.62	.80
	6	7.00	3.41	4.71	2.70	-.46	.75
	7	6.68	3.75	3.80	3.74	.51	2.51
2	4	5.48	2.73	3.21	2.23	-.96	.93
	5	6.91	3.18	4.68	1.80	-.76	1.48
	6	8.93	4.82	5.30	3.73	.08	1.61
	7	9.43	4.77	6.48	4.06	-.05	1.93
3	4	7.55	4.34	4.75	1.44	-1.36	.40
	5	9.55	4.82	6.21	3.39	-.72	2.06
	6	9.77	5.32	6.43	3.29	.81	2.36
	7	10.89	6.93	7.61	4.09	.15	2.36

Table 68

MANOVA Table for Hypothesis 5

F-Ratio for Multivariate Test
of Equality of Mean Vectors = 1.6304

Degrees of Freedom for Hypothesis = 18

Degrees of Freedom for Error = 1411.87

$p < .0459$

Table 68 (continued)

Variable	Between Mean Square	Univariate F	P Less Than
Test I	8.73	1.54	.1641
Test II	4.44	1.23	.2916
Test III	10.09	2.32	.0324

Degrees of Freedom for Hypothesis = 6

Degrees of Freedom for Error = 501

3 Covariates Have Been Eliminated

Hypothesis 6: There is no difference in the mean performances of boys and girls across the four grade levels.

The mean scores for the interaction sex x grade are given in Table 69. The F statistics for this interaction appears in Table 70. The results reported in Table 70 clearly indicate that there is no significant interaction of sex x grade. Hypothesis 6 can not be rejected.

Hypothesis 7: There is no difference in the mean performances of children in the three I.Q. groups across the two sexes and four grade levels.

The mean scores for the interaction I.Q. x sex x grade are reported in Table 71. In Table 72 appear the multivariate and univariate F statistics for this interaction. All F ratios are less than 1 indicating there is no significant interaction among the factors of I.Q., sex and grade. Thus hypothesis 7 can not be rejected.

Table 69

Mean Scores of Interaction, Sex x Grade

Sex	Grade	Raw Score Means Test I	Raw Score Means Test II	Raw Score Means Test III	Adjusted Means Test I	Adjusted Means Test II	Adjusted Means Test III
M	4	5.05	2.89	3.80	1.25	-.98	.77
	5	6.80	3.52	4.88	2.49	-.56	1.39
	6	8.65	4.53	5.44	3.01	.40	1.28
	7	8.89	5.02	6.29	3.78	.09	1.99
F	4	5.96	2.88	3.27	1.78	-1.07	.65
	5	7.56	3.11	4.52	2.46	-.82	1.51
	6	8.49	4.50	5.52	3.46	-.11	1.86
	7	9.11	5.29	5.64	4.24	.32	2.54

Table 70

MANOVA Table for Hypothesis 6

F-Ratio for Multivariate Test
of Equality of Mean Vectors = 1.2171

Degrees of Freedom for Hypothesis = 9

Degrees of Freedom for Error = 1214.59

$p < .2802$

Variable	Between Mean Square	Univariate F	P Less Than
Test I	3.92	.69	.5586
Test II	4.49	1.24	.2946
Test III	5.44	1.25	.2911

Degrees of Freedom for Hypothesis = 3

Degrees of Freedom for Error = 501

3 Covariates Have Been Eliminated

Table 71

Mean Scores of Interaction, I.Q. x Sex x Grade

I.Q.	Sex	Grade	Raw Score Means Test I	Raw Score Means Test II	Raw Score Means Test III	Adjusted Means Test I	Adjusted Means Test II	Adjusted Means Test III
1	M	4	3.05	1.46	2.64	.22	-1.07	.73
		5	4.86	2.14	3.09	2.58	-.63	1.62
		6	7.36	3.64	5.00	2.61	-.08	.37
		7	6.32	3.50	4.00	3.59	.14	2.02
	F	4	3.91	1.73	2.68	1.52	-.89	.89
		5	5.32	1.73	3.32	1.90	-.62	-.02
		6	6.64	3.18	4.41	2.78	-.83	1.13
		7	7.05	4.00	3.59	3.89	.88	2.99
2	M	4	4.73	2.64	3.68	2.59	-.72	1.29
		5	6.59	3.05	5.09	1.56	-.67	1.12
		6	8.82	4.41	5.00	3.67	.32	1.42
		7	9.59	4.27	7.05	3.82	.38	1.61
	F	4	6.23	2.82	2.73	1.87	-.76	.57
		5	7.23	3.32	4.27	2.03	-.85	1.85
		6	9.05	5.23	5.59	3.80	-.16	1.80
		7	9.27	5.27	5.91	4.56	-.47	2.25
3	M	4	7.36	4.59	5.09	.93	-1.17	.30
		5	8.96	5.36	6.46	3.35	-.46	1.42
		6	9.77	5.55	6.32	2.76	.95	2.06
		7	10.77	7.27	7.82	3.92	-.24	2.34
	F	4	7.73	4.09	4.41	1.94	-1.56	.50
		5	10.14	4.27	5.96	3.44	-.98	2.71
		6	9.77	5.09	6.55	3.81	.67	2.66
		7	11.00	6.59	7.41	4.25	.55	2.39

Table 72

MANOVA Table for Hypothesis 7

F-Ratio for Multivariate Test of Equality of Mean Vectors = .5374

Degrees of Freedom for Hypothesis = 18

Degrees of Freedom for Error = 1411.87

$p < .9414$

Variable	Between Mean Square	Univariate F	P Less Than
Test I	3.84	.68	.6697
Test II	0.97	.27	.9516
Test III	3.06	.70	.6480

Degrees of Freedom for Hypothesis = 6

Degrees of Freedom for Error = 501

3 Covariants Have Been Eliminated

Part 2: Correlation Studies

In the Statement of the Problem, Chapter II, a question was presented regarding the relationship of I.Q. with the performances on the probability tests. In addition, several questions were presented with regard to the relationship between the performance scores on the tests and subtests. Pearson product-moment correlation coefficients were computed to help answer these questions. This section presents the results of the correlation studies and a discussion of these results.

The Total I.Q. on the California Test of Mental Maturity was selected as a stratifying variable for this study. It was assumed that the Total I.Q. would be a better predictor of the performance on each of the probability tests than either the Language I.Q. or Non-Language I.Q. obtained from the California Test data. Question 8 was presented to test this assumption for the sample.

Question 8: Which of the three available scores on the California Test of Mental Maturity; Language I.Q., Non-Language I.Q., or Total I.Q. is the best predictor of the performance scores on the three probability tests?

Table 73 gives the correlations between the three I.Q.'s and the performance scores on each probability test for the four grades. An examination of this table shows that all but one of the correlations are significantly different ($p < .01$) from zero. These results also show that, except for Test II in grades six and seven, Total I.Q. has the highest correlations with the test scores in grades five, six and seven. In grade four Total I.Q. has the lowest correlations with the test scores. However, the differences between the correlations are very small and are not significant. In general, even though the differences are quite small, Total I.Q. appears to have been the best choice for the stratifying variable in this study.

Table 73

Correlations Between the Three Scores on the California Test of Mental Maturity and the Scores on the Three Probability Tests for the Children in the Four Grades

	Grade 4		
	Test I	Test II	Test III
Language I.Q.	.50**	.53**	.37**
Non-Language I.Q.	.47**	.53**	.39**
Total I.Q.	.37**	.40**	.25**
	Grade 5		
	Test I	Test II	Test III
Language I.Q.	.57**	.49**	.48**
Non-Language I.Q.	.54**	.48**	.48**
Total I.Q.	.66**	.58**	.56**
	Grade 6		
	Test I	Test II	Test III
Language I.Q.	.42**	.48**	.34**
Non-Language I.Q.	.37**	.16	.29**
Total I.Q.	.48**	.39**	.38**
	Grade 7		
	Test I	Test II	Test III
Language I.Q.	.56**	.58**	.52**
Non-Language I.Q.	.55**	.44**	.41**
Total I.Q.	.63**	.58**	.54**

** $p < .01$ * $p < .05$

Question 9: What is the relationship between the performance scores on the three probability tests within each grade?

The correlations between the performance scores on the three probability tests within each of the four grades appear in Tables 74, 75 and 76. All of the correlation coefficients are significantly different ($p < .01$) from zero.

The high correlations between performance scores on the three tests are not surprising because of the close association between the ideas included in the three tests. As reported in Chapter II, Test I dealt with the notion of sample space and Test II involved the idea of probability of a simple event. The items on Test III involved the notion of sample space and probability of a simple event. Each item in Test II presented a situation which was similar to the situation presented in the corresponding item in Test I. The question asked in each item on Test I was an implicit question in the corresponding item in Test II. That is, in order to answer a question in Test II, involving the probability of a simple event, the child had to understand what the sample space for the situation was. The items of Test I asked the subjects to list the elements of the sample space for each situation. Because of the similarity between situations in the corresponding items on Tests I and II one may expect that the correlations between performance scores on this pair of tests would be higher than the correlations between other pairs of tests. This was true for grades four and six but the differences are not significant.

To gain a better insight into the relationship between performance scores on Tests I and II several additional correlation studies between performance scores on the subtests were conducted.

Table 74

Correlations Between Scores on Test I and Test II

	Test II Grade 4	Grade 5	Grade 6	Grade 7
Test I	.58**	.47**	.50**	.48**

**p < .01

Table 75

Correlations Between Scores on Test I and III

	Test III Grade 4	Grade 5	Grade 6	Grade 7
Test I	.47**	.43**	.40**	.50**

**p < .01

Table 76

Correlations Between Scores on Test II and Test III

	Test III Grade 4	Grade 5	Grade 6	Grade 7
Test II	.42**	.49**	.40**	.49**

**p < .01

Question 10: What is the relationship between performance scores on Subtest I-A and Subtest II-A within each grade?

Table 77 presents the mean scores for Subtests I-A and II-A. The correlations between the total scores on these subtests for each grade are given in Table 78. Each of these correlations is significantly different from zero. Because of the reasonably high mean scores on these subtests and the similarities in the basic situations presented in corresponding items of the subtests one would suspect that these correlations would be higher. It is difficult to understand how a child can answer a question about the probability of a simple event correctly and yet not be able to list all of the elements of the sample space which contains this event. However, many children made this type of error which partially accounts for the lower than expected correlations.

Table 79 contains the correlations for the six pairs of corresponding items on Subtests I-A and II-A for each grade.[71] The very low correlations between these items are very surprising. About one-third of the correlations are not significantly different from zero. The only pairs of items which are significantly correlated within each grade are items 5 and 17 and items 6 and 18. These items involve sampling without replacement. These items were also the most difficult items on each of the subtests. Apparently the idea of sampling without replacement not only made the items more difficult but also caused these pairs of corresponding items on the subtests to function more dependently.

The fact that the other pairs of corresponding items functioned independently in most cases can probably be explained in part by the nature of the study. The subjects had no formal training on the ideas of probability and therefore had not been taught the relationship between sample space and probability of a simple event. It is reasonable to assume that if these tests were administered after the children had received some instruction on the underlying concepts, the performance scores would improve and the correlations between corresponding items would be substantially higher.

[71]The complete inter-item correlation matrices for Subtests I-A and II-A are reported in Appendix B.

Table 77

Mean Scores for Subtest I-A and Subtest II-A

Grade	Subtest I-A Mean	Subtest I-A Per Cent Correct	Subtest II-A Mean	Subtest II-A Per Cent Correct
4	3.56	59%	2.28	38%
5	4.20	70%	2.65	44%
6	4.77	78%	3.55	59%
7	5.00	83%	3.97	61%

Table 78

Correlations Between Total Scores on Subtest I-A and Subtest II-A

	Subtest I-A Grade 4	Grade 5	Grade 6	Grade 7
Subtest II-A	.42**	.25**	.43**	.47**

**p < .01

Table 79

Correlations Between Corresponding Items on Subtests I-A and II-A

Items	Grades 4	5	6	7
1 and 13	.13	.07	.19*	.30**
2 and 14	.19*	.05	.13	.18*
3 and 15	.22**	.08	.14	.15
4 and 16	.10	.12	.25**	.20*
5 and 17	.29**	.19*	.30**	.37**
6 and 18	.39**	.29*	.36**	.30**

*p < .05 **p < .01

Question 11: What is the relationship between the performance scores on Subtest I-B and Subtest II-B within each grade?

The means scores for Subtests I-B and II-B appear in Table 80. The correlations between total scores on the subtests are given in Table 81. The correlations between total scores are significantly different from zero but are very small. These low correlations are undoubtably due in part to the fact that Subtest II-B was extremely difficult for all grades. An examination of Table 80 indicates that children in the fifth, sixth and seventh grades exhibit some understanding of combinations in situations involving the idea of sample space but are not able to apply these ideas in similar situations involving probability of a simple event.

Table 82 give the correlations between the six pairs of corresponding items on Subtests I-B and II-B.[72] All of the correlations reported in this table are very small. The majority of coefficients are not significantly different from zero. Five of the correlations are significant at the .05 level.

---

[72]The complete inter-item correlation matrices for Subtests I-B and II-B are reported in Appendix C.

Table 80

Mean Scores for Subtest I-B and Subtest II-B

Grade	Subtest I-B Mean	Subtest I-B Per Cent Correct	Subtest II-B Mean	Subtest II-B Per Cent Correct
4	1.92	32%	.62	10%
5	2.97	49%	.66	11%
6	3.80	63%	.97	16%
7	4.00	67%	1.18	19%

Table 81

Correlations Between Total Scores on Subtest I-B and Subtest II-B

	Subtest I-B			
	Grade 4	Grade 5	Grade 6	Grade 7
Subtest II-B	.27**	.32**	.29**	.27**

** p < .01

Table 82

Correlations Between Corresponding Items on Subtests I-B and II-B

	Grades			
Items	4	5	6	7
7 and 19	.13	.19*	.03	.04
8 and 20	.11	.10	.05	.16
9 and 21	.16	.16	.04	.11
10 and 22	.00	.09	.04	.12
11 and 23	.16	.16	.10	.20*
12 and 24	.20*	.18*	.12	.18*

* p < .05 ** p < .01

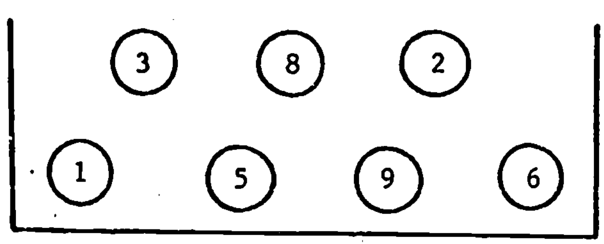
Part 3: Analysis of Incorrect Responses on the Test Items

In Chapter IV item statistics, which included an item difficulty index, inter-item criterion correlation, X_{50} and β for each item, were reported for each test based on the performances of children in each grade. These statistics helped explain how each item functioned at the different grade levels but did not provide information which would help explain why some items were more difficult than others. This part includes a summary and analysis of the incorrect responses on each item. The patterns of errors on each item were examined to gain some insight into the misconceptions children may have. Since the tests were administered as written tests, and children were not interviewed, one can only make conjectures about how children thought about the items. However, the patterns of errors on certain items leave little doubt about the pattern of thought employed.

The analyses of errors for the test items are presented individually on the following pages. For the convenience of the reader the per cent of children in each grade who answered the item incorrectly is included under the statement of the item. The reader is asked to refer to the tables in Chapter IV for the other item statistics.

Subtest I-A

1. For this experiment a box contains balls as in the picture. To do this experiment you pick one ball from the box without looking. The number that is on the ball that you pick is called an out-come of this experiment.



In the space below, write all the different outcomes it would be possible to obtain for this experiment.

Table 83

Per Cent of Incorrect Responses on Item 1

Grade 4	Grade 5	Grade 6	Grade 7
33%	19%	6%	8%

This item was marked wrong if the child did not list all seven of the possible outcomes or if he listed more than seven outcomes.

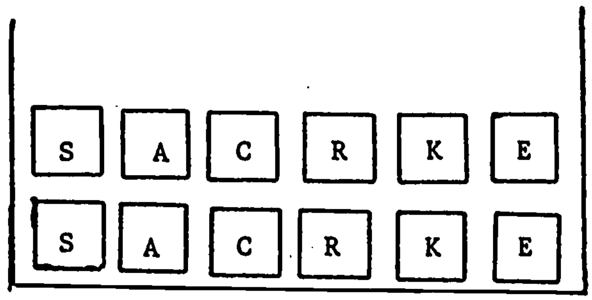
The item was very easy for grades six and seven. Although more errors were made by children in grades four and five it can be considered an easy item for all grades.

All of the errors on this item were due to not listing all seven of the possible outcomes for the experiment described in the item. No one listed more than seven outcomes.

Since this was the first item on the test it is presumed that some of the errors were due to a lack of understanding of what was expected. This was particularly evident in Grade 4 where approximately 10% of the children that had the item wrong did not write anything on their papers. On item 2 only 1% of the children failed to write an answer on their papers. It appears that some children needed a little more direction than was given in the two sample items.

It is worth noting that only 10 of the children that missed item 1 did not make any other errors on the first six items. The majority of children who made an error on item 1 had three or more errors on Subtest I-A.

2. For this experiment a box contains cards as in the picture. To do this experiment you pick one card from the box without looking. The letter that is on the card that you pick is called an outcome of this experiment.



In the space below, write all the different outcomes it would be possible to obtain for this experiment.

Table 84

Per Cent of Incorrect Responses on Item 2

Grade 4	Grade 5	Grade 6	Grade 7
29%	29%	28%	20%

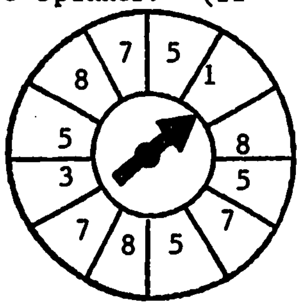
This item was marked wrong if the child listed fewer or more than the six different outcomes for the experiment.

This item can also be considered an easy item for all grades. It is interesting to note that approximately the same number of children in each grade made an error on this item.

The majority of errors on this item were due to listing more than the six different outcomes for the experiment. Many children simply listed all of the six letters twice. This indicates a misunderstanding of the meaning of the word "different" as it is used in the context of the item.

Since the meaning of the word "different" was emphasized in the sample items, and the word was again given special attention when the item was read aloud, it was decided to mark the item wrong if the child repeated any of the letters in his list of outcomes. Approximately 50% of the children who repeated letters in their lists only repeated some of the letters, not all six. This suggests that something else is involved besides a misunderstanding of the meaning of the word "different." It is not clear why children made this type of error.

3. For this experiment a spinner is marked as in the picture. To do this experiment you spin the arrow on the spinner. (If the arrow stops on a line you spin it again.) The number that the arrow points to when it stops is called an outcome of this experiment.



In the space below, write all the different outcomes it would be possible to obtain for this experiment.

Table 85

Per Cent of Incorrect Responses on Item 3

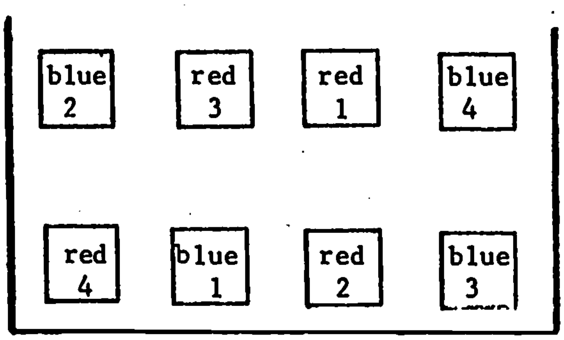
Grade 4	Grade 5	Grade 6	Grade 7
32%	29%	27%	21%

This item was marked wrong if the child listed fewer or more than the five different outcomes for this experiment.

This item is very similar to item 2. The per cent of errors for each grade on this item is the same or very close to the per cent of errors on item 2. However, a considerable number of children (34%) who responded incorrectly on this item had item 2 correct. The pattern of errors on this item is somewhat different than item 2. The majority of errors were due to listing some of the outcomes more than once but very few children listed all of the numerals on the figure. For example, many children listed 5 more than once but very few listed 5 four times although 5 appears in four different places on the spinner. This type of error again suggests a possible misunderstanding of the meaning of the word "different." A few children failed to include 3 in their list but did list 5, 7, 1 and 8. This type of error suggests that these children did not study the figure carefully but appear to understand the idea of sample space in this situation.

4. For this experiment a box contains cards as in the picture.

To do this experiment you pick one card from the box without looking. The <u>color-number pair</u> that is on the card that you pick is called an <u>outcome</u> of this experiment. For example, one outcome is the color-number pair (red, 3).



In the space below, write <u>all the different</u> outcomes it would be <u>possible to obtain</u> for this experiment.

Table 86

Per Cent of Incorrect Responses on Item 4

Grade 4	Grade 5	Grade 6	Grade 7
42%	21%	14%	11%

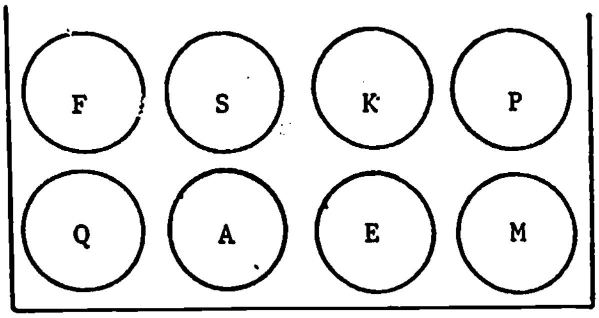
This item was marked wrong if the child listed fewer or more than the eight different outcomes for this experiment.

As indicated by the results in Table 86 this item was quite easy for all grades except grade four.

In grades six and seven the majority of errors were due to listing fewer or more than the eight outcomes. Many of the children in grades four and five also made this type of error. The common error was to list only four pairs with the numerals 1, 2, 3 and 4 used only once or to list two pairs with the color words red and blue used only once. These children failed to recognize that the pair (blue, 2) was different than the pair (red, 2) and so on. Those few children who listed more than eight outcomes, listed pairs such as (red, blue), (1,2), etc., which indicated they did not understand the basic idea presented in the item. This latter type of error was very rare.

Another type of error exhibited by a few children in grades four and five was a list of only numerals or just the two color words. These children did not understand the notion of a color-number pair even though a sample was included in the item.

5. For this experiment a box contains chips as in the picture.



To do this experiment you pick one chip from the box without looking. The letter that is on the chip that you pick is called an outcome of this experiment.

Imagine that the first chip you pick as a "K" on it. You do not put this chip back into the box. Then you pick a second chip.

In the space below, write all the different outcomes it would be possible to obtain on the second pick.

Table 87

Per Cent of Incorrect Responses on Item 5

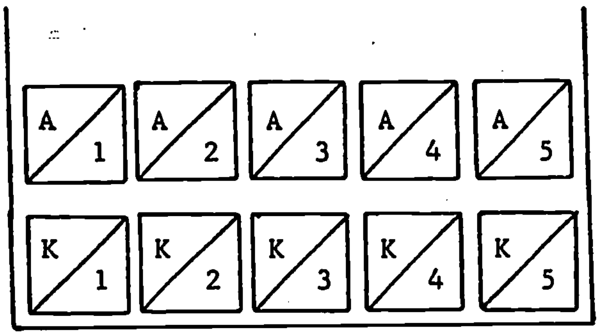
Grade 4	Grade 5	Grade 6	Grade 7
50%	35%	19%	17%

This item was marked wrong if the child listed fewer or more than the seven different outcomes for the experiment.

This item was considerably more difficult for fourth and fifth grade children than for sixth and seventh grade children.

More than 80% of the errors on this item in each grade were due to including K as a possible outcome on the second pick. This clearly indicates that these children did not understand the idea of sampling without replacement in this experiment. The other errors were due to listing fewer than seven outcomes with answers ranging from one to six outcomes. These children did not have a good notion of sample space or may have misunderstood the meaning of the word "different."

6. For this experiment a box contains cards as in the picture. To do this experiment you pick one card from the box without looking. The letter/number pair that is on the card that you pick is called an outcome of this experiment. For example, one outcome is the letter/number pair (K/4).



Imagine that the first two cards that you pick have the number "4" on them. You do not put these cards back into the box. Then you pick a third card.

In the space below, write all the different outcomes (letter/number pairs) that it would be possible to obtain on the third pick.

Table 88

Per Cent of Incorrect Responses on Item 6

Grade 4	Grade 5	Grade 6	Grade 7
58%	47%	30%	22%

The item was marked wrong if the child listed fewer or more than the eight different outcomes for this experiment.

Item 6 is very similar to item 5 and functioned in much the same way. The two items are highly correlated with inter-item correlations of .67, .65, .66 and .63 for grades four through seven respectively. Item 6 was more difficult than item 5 because it involved not replacing two of the objects in the box and the outcomes were described as pairs rather than as a single letter or numeral.

The majority of errors were due to listing one or both of the pairs A/4 and K/4 as possible outcomes on the third pick. This indicates a lack of understanding of sampling without replacement and in this respect is very similar to item 5. However this type of error was not as prevalent on item 6 as on item 5. The other errors were due mainly to listing fewer than eight outcomes but in many cases the few outcomes that were listed included the pair A/4 or the pair K/4. A few of the fourth grade children did not understand the idea of the letter/number pair and they listed only letters or only numerals.

Summary of Errors on Subtest I-A

From the responses on Subtest I-A it is apparent that:

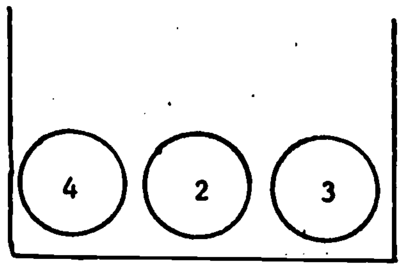
a) Some children did not understand the meaning of the word "different" as it was used in the context of the items.

b) Children in the fourth grade have some difficulty in listing pairs as outcomes.

c) The notion of sampling without replacement is a very difficult idea and children tended to keep the object that has been removed as part of the sample space.

Subtest I-B

7. For this experiment a box contains chips as in the picture. To do this experiment you pick two chips from the box at the same time without looking. The sum of the numbers on the two chips that you pick is called an outcome of this experiment. For example, one outcome is the sum (4 + 2) or 6.



In the space below, write all the different outcomes (sums) it would be possible to obtain for this experiment.

Table 89

Per Cent of Incorrect Responses on Item 7

Grade 4	Grade 5	Grade 6	Grade 7
44%	33%	19%	23%

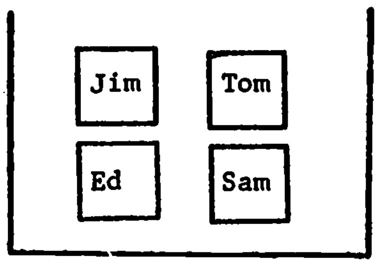
This item was marked wrong if the child listed fewer or more than the three different outcomes for the experiment.

The pattern of errors was very similar in all grades even though the item was somewhat more difficult for fourth and fifth grade children. Of the items marked wrong, about 30% did not include any of the outcomes, about 20% had one of the three outcomes and the other 50% had two of the three outcomes.

Most of the children who listed two of the three combinations, listed 4 + 2 and 2 + 3 but missed 4 + 3. It is clear that the position of the chips in the diagram helped the children recognize the combinations 4 + 2 and 2 + 3.

8. For this experiment a box contains slips with names on them as in the picture.

To do this experiment you pick <u>two</u> slips from the box at the same time without looking.



The <u>pair of names</u> on the two slips that you pick is called an <u>outcome</u> of this experiment. For example, one outcome is the pair of names (Ed, Sam).

In the space below, write <u>all the different outcomes</u> (pairs of names) it would be <u>possible to obtain</u> for this experiment.

Table 90

Per Cent of Incorrect Responses on Item 8

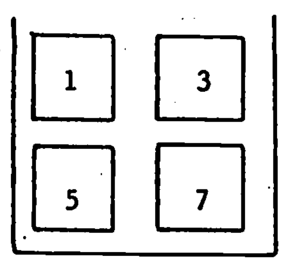
Grade 4	Grade 5	Grade 6	Grade 7
69%	51%	33%	32%

This item was marked wrong if the child listed fewer or more than the six different combinations for this experiment.

About 60% of the sixth and seventh grade children for whom this item was marked wrong had listed four or five of the six pairs of names. In general, it appeared like they missed some of the pairs because they did not list the pairs systematically.

The fourth and fifth grade children who made an error on this item did not have a very good idea about combinations. About 30% of these children did not list any of the pairs and another 30% listed only two pairs. The two pairs usually were (Ed, Sam) and (Jim, Tom) which are the obvious pairs suggested by the arrangement of names in the diagram.

9. For this experiment a box contains cards as in the picture. To do this experiment you pick <u>three</u> cards from the box at the same time without looking.



The sum of the numbers on the three cards that you pick is called an <u>outcome</u> of this experiment. For example, one outcome is the sum (1 + 3 + 7) or 11.

In the space below, write <u>all the different outcomes</u> (sums) it would be <u>possible to obtain</u> for this experiment.

Table 91

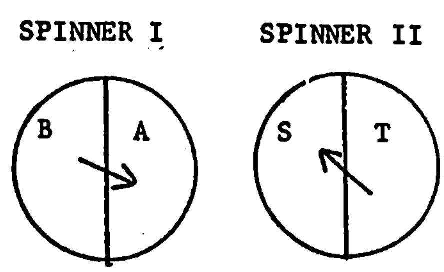
Per Cent of Incorrect Responses on Item 9

Grade 4	Grade 5	Grade 6	Grade 7
74%	59%	49%	51%

This item was marked wrong if the child listed fewer or more than the four combinations for this experiment.

The pattern of errors was very similar for all grades. Approximately 48% of the children who made an error on this item did not list any outcomes or listed only the sample outcome. About 15% listed only two outcomes and the other 37% listed three of the four outcomes. The children who listed three of the four combinations again demonstrated some understanding of combinations but in general did not list the outcomes systematically.

10. For this experiment two spinners are marked as in the picture. To do this experiment you spin the arrow on each spinner. (If an arrow stops on a line you spin it again.) The pair of letters in the spaces that the two arrows point to when they stop is called an outcome of this experiment.



For example, one outcome is the pair of letters (A, S).
In the space below, write all the different outcomes (pairs of letters) it would be possible to obtain for this experiment.

Table 92

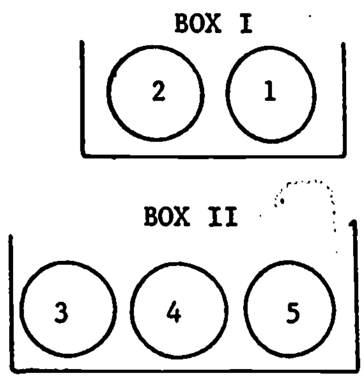
Per Cent of Incorrect Responses on Item 10

Grade 4	Grade 5	Grade 6	Grade 7
71%	55%	44%	38%

This item was marked wrong if the child listed fewer or more than the four combinations for this experiment.

Apparently the interpretation of the phrase "on each spinner" caused some difficulty. About 65% of the errors were due to listing pairs like (B,A), (B,B), (S,T) and so on, or not listing any of the pairs. Another 25% of the errors were due to children listing only two of the pairs. In general the two pairs were (A,S) and (B,T).

11. For this experiment two boxes contain balls as in the picture. To do this experiment you pick one ball from each box without looking. The product of the numbers on the two balls that you pick is called an outcome of this experiment. For example, one outcome is the product (2 x 4) or 8.



In the space below, write all the different outcomes (products) it would be possible to obtain for this experiment.

Table 93

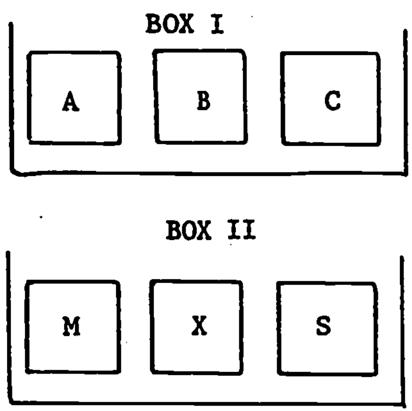
Per Cent of Incorrect Responses on Item 11

Grade 4	Grade 5	Grade 6	Grade 7
76%	53%	38%	28%

This item was marked wrong if the child listed fewer or more than the six combinations for this experiment.

Some children had difficulty with the interpretation of the phrase "from each box." A large majority of the errors were due to children listing products like (2 x 2), (3 x 4), and so on. Approximately 80% of the errors on this item were due to listing products like the above or not listing any products. The other errors were due to incomplete lists which generally reflected a lack of a systematic approach for obtaining all possible combinations.

12. For this experiment two boxes contain cards as in the picture. To do this experiment you pick one card from each box without looking. The pair of letters on the two cards that you pick is called an outcome of this experiment. For example, one outcome is the pair of letters (C,X).



In the space below, write all the different outcomes (pairs of letters) it would be possible to obtain for this experiment.

Table 94

Per Cent of Incorrect Responses on Item 12

Grade 4	Grade 5	Grade 6	Grade 7
74%	53%	38%	29%

This item was marked wrong if the child listed fewer or more than the nine combinations for this experiment.

This item is very similar to item 11 and functioned in much the same way. The results reported in Tables 93 and 94 are almost identical and the items are highly correlated. The inter-item correlations for items 11 and 12 are .61, .78, .87 and .68 for grades four through seven respectively. About 65% of the errors were due to listing pairs like (A,A) and (K,S) or not listing any of the pairs. The other errors were due to incomplete lists and again reflected a lack of a system for obtaining all possible pairs.

Summary of Errors on Subtest I-B

From the responses on Subtest I-B it is apparent that:

a) Some children used the position of the objects in the figure for obtaining the combinations and did not consider the less obvious combinations.

b) The meaning of the word "each" in a phrase like "from each box" was misinterpreted by many children.

c) Many children have not developed a systematic method for generating combinations and consequently often omitted one or two of the combinations in an item.

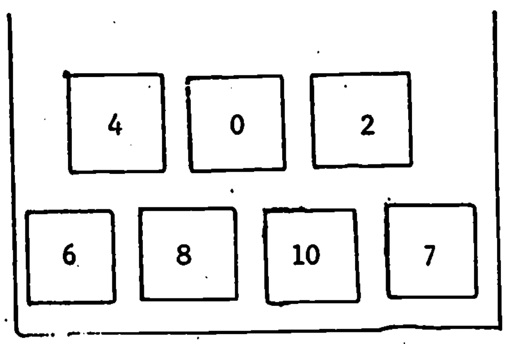
Subtest II-A

13. A box contains cards as in the picture.

To play this game you pick one card from the box without looking.

You win if you pick the card with the "2" on it.

You lose if you pick a card with any number on it.



If you play this game only once, what chance do you have of winning?

Answer: ____ out of ____

Table 95

Per Cent of Incorrect Responses on Item 13

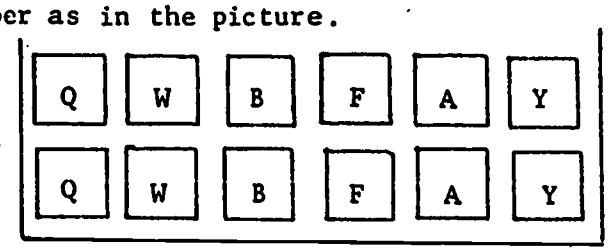
Grade 4	Grade 5	Grade 6	Grade 7
36%	31%	15%	17%

This item was marked wrong if the child did not give the response "1 out of 7."

Approximately 45% of the wrong answers were given as "1 out of 6." This gives the odds for winning the game rather than the probability or chance of winning. It was decided to mark this answer wrong even though it indicates that children who give this type of answer may have some notions of probability. Examining the pattern of responses for all the items on Test II showed that children who used the "odds" representation were not consistent in this usage. This inconsistency gave substance to the decision to consider the "odds" response as an error.

The other errors were extremely varied and did not suggest any particular patterns. The error "2 out of 7" did appear on a number of papers. This error may be due to the children confusing the symbol "2" on the winning card with the chance of picking the winning card.

14. A box contains slips of paper as in the picture. To play this game you pick one slip of paper from the box without looking.



<u>You win if you pick a slip with an "A" on it</u>.

You lose if you pick a slip with any other letter on it.

If you play this game only once, what chance do you have of winning?

Answer: _____ out of _____

Table 96

Per Cent of Incorrect Responses on Item 14

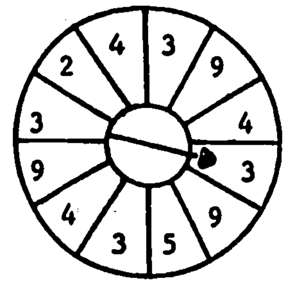
Grade 4	Grade 5	Grade 6	Grade 7
67%	53%	38%	27%

This item was marked wrong if the child did not give the response "2 out of 12" or an equivalent response such as "1 out of 6."

The majority of errors were of two types. Approximately 38% of the children who made an error gave the answer "2 out of 10." This represents the odds for winning the game, not the probability of winning. Another 40% of the wrong answers were given as "1 out of 12." These children obviously did not consider both of the slips with "A" on them as winners. This type of error may be due to children thinking that they pick one card from a box of twelve cards so the chance of winning must be 1 out of 12, disregarding the number of cards which are potential winners.

The other errors were varied and did not show any particular patterns.

15. A spinner is marked as in the picture. To play this game you spin the arrow on the spinner. (If the arrow stops on a line you spin it again.)



<u>You win if the arrow points to a space marked with a "4" when it stops</u>.

You lose if the arrow points to a space with any other number on it.

If you play this game only once, what chance do you have of winning?

Answer: _____ out of _____

Table 97

Per Cent of Incorrect Responses on Item 15

Grade 4	Grade 5	Grade 6	Grade 7
70%	58%	43%	35%

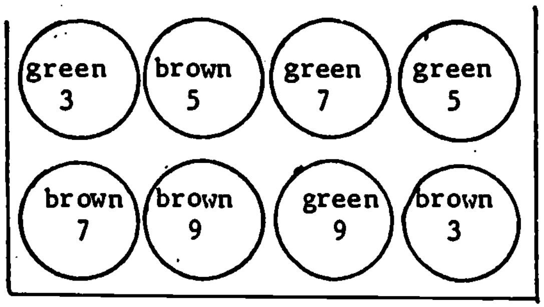
This item was marked wrong if the child did not give the response "3 out of 12" or "1 out of 4."

This item is very similar to item 14 and functioned in much the same way in all grades. Items 14 and 15 are highly correlated with inter-item correlations of .58, .58, .62 and .55 for grades four through seven respectively.

The pattern of errors for item 15 is also similar to the pattern for item 14. Forty-three per cent of the children who made an error on this item gave the odds of winning, "3 out of 9," as their response.

The response "1 out of 12" accounted for approximately 30% of the errors. The other errors did not suggest any particular patterns.

16. A box contains balls as in the picture.



To play this game you pick one ball from the box without looking.

<u>You win if you pick a green ball with a "5" on it.</u>

You lose if you pick any other ball.

If you play this game only once, what chance do you have of winning?

Answer: _____ out of _____

Table 98

Per Cent of Incorrect Responses on Item 16

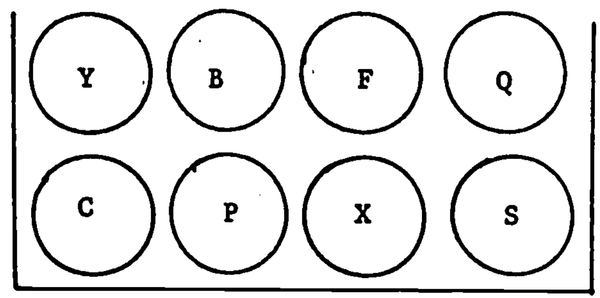
Grade 4	Grade 5	Grade 6	Grade 7
46%	42%	36%	24%

This item was marked wrong if the child did not give the response "1 out of 8."

This item is similar to item 1 but somewhat more difficult. Part of the difficulty was probably due to the fact that the balls in the game described in this item are marked with both a color and a numeral. The balls in item 1 are marked with only a numeral. The duplication of colors and numerals also accounted for some of the difficulty.

Of the errors, 23% were given as "1 out of 7," the odds for winning the game. Approximately 30% of the wrong answers were "1 out of 8" and another 10% of the errors were "4 out of 8." These errors can probably be explained by the duplication of two colors and four numerals on the balls in the figure.

17. A box contains chips as in the picture.
To play this game you pick one chip from the box without looking.
You win if you pick the chip with the "X" on it.
You lose if you pick a chip with any other letter on it.



Imagine that the first chip you pick has "B" on it and is not a winner. You do not put this chip back into the box. Then you pick a second chip. What chance do you have of winning on the second try?

Answer: _____ out of _____

Table 99

Per Cent of Incorrect Responses on Item 17

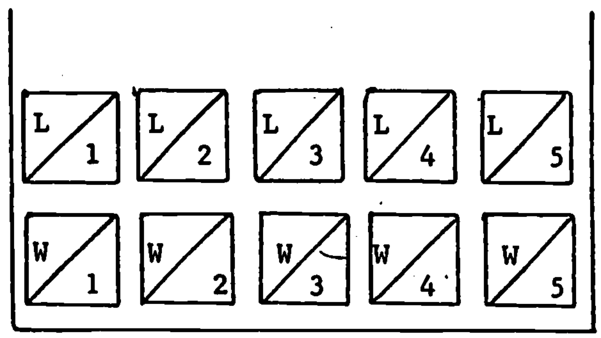
Grade 4	Grade 5	Grade 6	Grade 7
67%	67%	46%	40%

This item was marked wrong if the child did not give the response "1 out of 7."

This item involved sampling without replacement which undoubtably accounted for many of the errors. This is evident from the fact that 43% of the errors were given as "1 out of 8." Only 8% of the errors on this item gave the odds for winning, "1 out of 6." About 16% of the errors were given as "2 out of 8." The later error was probably due to the children thinking of 2 balls picked from a box containing 8 balls.

18. A box contains cards as in the picture.

To play this game you pick one card from the box without looking.
<u>You win if you pick a card with a "W" on it.</u>
You lose if you pick a card with a "L" on it.



Imagine that the first <u>two</u> cards that you pick have "L" on them and are not winners. You <u>do not</u> put these cards back into the box. Then you pick a <u>third</u> card.

What chance do you have of winning on the <u>third</u> try?

Answer: _____ out of _____

Table 100

Per Cent of Incorrect Responses on Item 18

Grade 4	Grade 5	Grade 6	Grade 7
87%	84%	67%	60%

This item was marked wrong if the child did not give the response "5 out of 8."

As can be seen from the results reported in Table 100 this was a very difficult item, particularly for the fourth and fifth grades.

Only 8% of the wrong answers gave the odds for winning, "5 out of 3." The other errors did not suggest any particular patterns except that the majority of incorrect responses had either 5 or 10 as the second numeral in the answer. As in item 17, much of the difficulty with the item was most likely due to the children's lack of understanding of sampling without replacement.

Summary of Errors on Subtest II-A

From the responses on Subtest II-A it is apparent that:

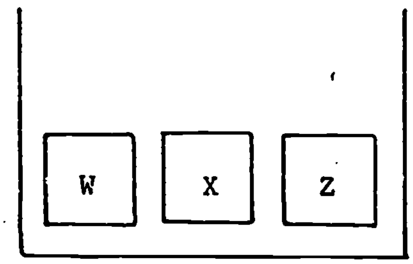
a) One cause for error was a lack of understanding of the idea of sample space. The inability to recognize all of the possible outcomes of a sample space lead to an incorrect statement of the probability of a simple event in the space.

b) Many children gave the odds for winning a game rather than the probability or chance of winning. These children tended to be inconsistent in their use of "odds" and "probability."

c) Some children gave responses such as "A out of 6," "4 out of 0," "6 out of K," and so on which suggests that these children have not yet acquired an understanding of probability of a simple event as it was used in the context of the items on Test II.

Subtest II-B

19. A box contains cards as in the picture.
To play this game you pick two cards from the box at the same time without looking.
You win if one of the two cards that you pick is the card with the "X" on it.
You lose if you do not pick the card with the "X" on it.



If you play this game only once, what chance do you have of winning?

Answer:_____ out of _____

Table 101

Per Cent of Incorrect Responses on Item 19

Grade 4	Grade 5	Grade 6	Grade 7
79%	79%	70%	67%

This item was marked wrong if the child did not give the response "2 out of 3."

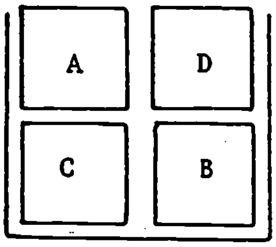
The great majority of errors on this item were of two forms. Approximately 70% of the children who made an error on this item gave the response "1 out of 3." These children undoubtably based their answer on the number of cards in the figure. There are three cards in the box and only one card has a "X" on it. Another 19% of the errors were the response "1 out of 2." No doubt this error was also due to the child's improper use of the figure. There is one card with "X" on it and two cards without "X" on them. These children most likely did not consider the possible combinations of two cards. Only 2% of the errors gave the odds for winning, "2 out of 1."

20. A box contains cards as in the picture.

To play this game you pick <u>two</u> cards from the box at the same time without looking.

<u>You win if you pick the pair of cards with "A" on one card and "B" on the other card</u>.

You lose if you do not pick this pair of cards.



If you play this game only once, what chance do you have of winning?

Answer: _____ out of _____

Table 102

Per Cent of Incorrect Responses on Item 20

Grade 4	Grade 5	Grade 6	Grade 7
99%	98%	98%	92%

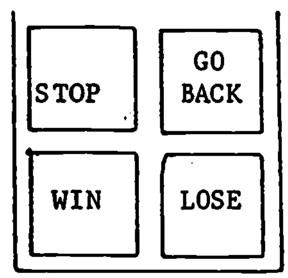
This item was marked wrong if the child did not give the response "1 out of 6."

As is evident from Table 102 this was an extremely difficult item and in fact was the most difficult of the 34 items on the tests. An examination of the errors clearly indicates that most children did not think about all the possible combinations of two letters but used the number of objects in the figure as the basis for their responses. The response "2 out of 4" accounted for 54% of the errors and another 23% of the errors were given as "1 out of 4." Less than 1% of the errors gave the odds for winning, "1 out of 5."

It is interesting to note that more than 50% of the subjects did list all of the six combinations of four things taken two at a time on the corresponding item on Test I. (See Table 91 for item 8).

21. A box contains slips of paper as in the picture.

To play this game you pick three slips from the box at the same time without looking. You win if one of the three slips that you pick has the word "WIN" on it. You lose if you do not pick the slip with the word "WIN" on it.



If you play this game only once, what chance do you have of winning?

Answer: _____ out of _____

Table 103

Per Cent of Incorrect Responses on Item 21

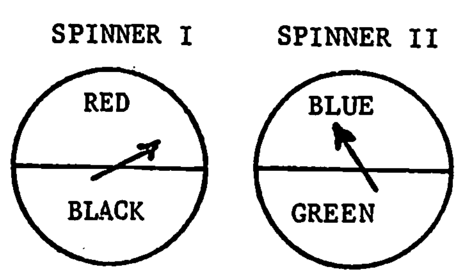
Grade 4	Grade 5	Grade 6	Grade 7
84%	80%	67%	65%

This item was marked wrong if the child did not give the response "3 out of 4."

Of the children who made errors on this item, 65% gave the response "2 out of 4" and 23% gave "1 out of 4" as their answer. About 2% of the errors were given as "3 out of 1." Since there are four possible combinations of four things taken three at a time and four slips in the box in the diagram one can not be certain how children thought about this problem. From the patterns of errors on the other items involving combinations the number of objects in the box was probably used as the basis for many of the answers.

It is suspected that some of the children who gave the correct answer for this item did not actually understand the item. The fact that there are four slips in the box and the directions for the game say to pick three of the slips may have prompted the response "3 out of 4." Therefore this is a very poor item for one can not be certain how the child arrived at his answer.

22. Two spinners are marked as in the picture. To play this game you spin the arrow on each of the spinners. (If an arrow stops on a line you spin it again.)



You win if the arrow on the first spinner points to a space marked red and the arrow on the second spinner points to a space marked blue when they stop. You lose if the arrows stop in any other way.

If you play this game only once, what chance do you have of winning?

Answer: _____out of_____

Table 104

Per Cent of Incorrect Responses on Item 22

Grade 4	Grade 5	Grade 6	Grade 7
80%	84%	78%	76%

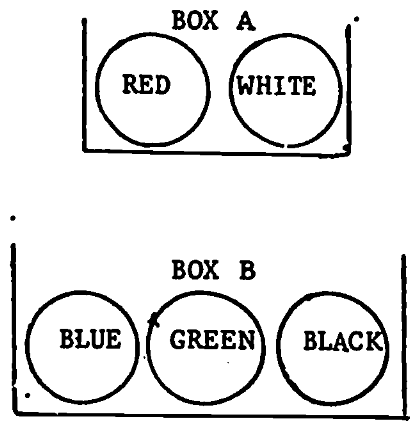
This item was marked wrong if the child did not give the response "1 out of 4."

Approximately 85% of the errors on this item were given as "2 out of 4." This type of error may be due to the fact that there are four spaces marked on the spinners and two of the spaces are marked red and blue. Another 5% of the errors were given as "2 out of 2."

Just as in item 21 one can not be certain how the child decided why the numeral 4 should be used as the second numeral in the response. He may have known that there are four combinations possible or he may have used the number of spaces on the spinners as the basis for this decision.

23. Box A and Box B contain colored chips as in the picture.

To play this game you pick one chip from each box without looking. You win if you pick the red chip from Box A and the blue chip from Box B. You lose if you do not pick this pair of chips.



If you play this game only once, what chance do you have of winning?

Answer: _____ out of _____

Table 105

Per Cent of Incorrect Responses on Item 23

Grade 4	Grade 5	Grade 6	Grade 7
98%	96%	95%	90%

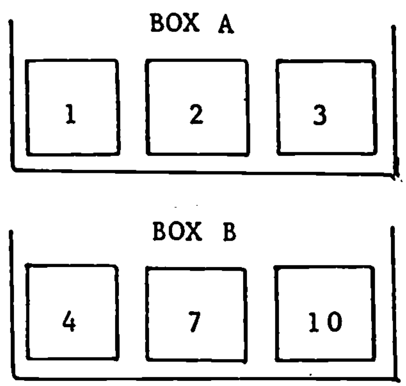
This item was marked wrong if the child did not give the response "1 out of 6."

This item was very difficult for all grades. The pattern of errors was similar for all grades and clearly indicates that many children based their answers on the number of chips in the two boxes. Approximately 64% of the wrong answers were given as "2 out of 5" and another 25% of the errors were given as "1 out of 5." The numeral 5 certainly was used because there are five balls in the two boxes.

These children did not consider the combinations even though many children (see Table 94) were able to list the six combinations for the corresponding item in Test 1.

24. Box A and Box B contain cards as in the picture.

To play this game you pick one card from each box without looking. You win if the sum of the numbers on the two cards is 6. You lose if the sum is any other number.



If you play this game only once, what chance do you have of winning?

Answer: _____ out of _____

Table 106

Per Cent of Incorrect Responses on Item 24

Grade 4	Grade 5	Grade 6	Grade 7
98%	97%	95%	90%

This item was marked wrong if the child did not give the response "1 out of 9."

Item 24 is similar to item 23 and functioned in much the same way. Notice that the results reported in Table 106 are almost identical to the results in Table 105. The inter-item correlations for items 23 and 24 are .32, .49, .49 and .62 for grades four through seven respectively.

Thirty-five per cent of the children who had this item wrong gave the answer "2 out of 6" and another 32% of the wrong answers were given as "1 out of 6." Obviously these children were influenced by the number of objects in the boxes and they did not compute all possible combinations. The results given in Table 94 show that subjects were able to list all nine of the combinations for the corresponding item on Test 1.

Summary of Errors on Subtest II-B

From the responses on Subtest II-B it is apparent that:

a) Many children used the number of objects in the figure for the item as the basis for their answer and did not consider the set of all possible combinations.

b) Some children lack an understanding of the notion of probability of a simple event for they gave responses such as "A out of 5," "2 out of X," and so on.

c) Difficulties such as confusion between "odds" and "probability," failure to list all possible combinations, and misinterpretation of a phrase like "from each box," accounted for many of the errors on these items.

Test III

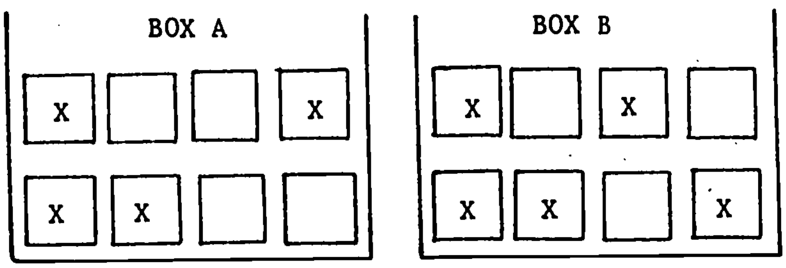
The summary and analyses of errors on the items of Test III are presented on the following pages. Since these items are multiple choice items with three options, the per cent of children in each grade who selected each option is presented under the statement for each item.

Since some pupils did not answer all of these items the sum of the per cents for the three options may be less than 100% on some items.

25. Box A and Box B are used to play this game. The boxes contain cards as in the picture.

To play this game you pick a card from one of the boxes without looking. You win if you pick a card with a "X" on it. You lose if you pick a blank card.

If you can play this game only once, which box would you choose to pick from so that you would have the better chance of winning?



_____ Box A

_____ Box B

_____ It doesn't make any difference

Table 107

Per Cent of Children Who Chose Each Option on Item 25

	Grade 4	Grade 5	Grade 6	Grade 7
Box A	10%	10%	7%	5%
Box B	77%	87%	89%	90%
No difference	13%	3%	4%	5%

Box B is the correct response for item 25.

As can be seen from the results in Table 107, this was a very easy item for all grades.

Some children may have selected the third option because there are eight cards in each box. It is not clear why some children picked Box A rather than one of the other options.

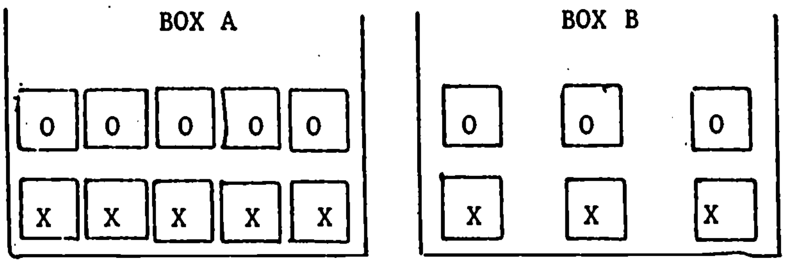
26. Box A and Box B are used to play this game. The boxes contain cards as in the picture.

To play this game you pick a card from one of the boxes without looking.

<u>You win if you pick a card with an "O" on it.</u>

You lose if you pick a card with an "X" on it.

If you can play this game only once, which box would you choose to pick from so that you would have the better chance of winning?



_____ Box A

_____ Box B

_____ It doesn't make any difference

Table 108

Per Cent of Children Who Chose Each Option on Item 26

	Grade 4	Grade 5	Grade 6	Grade 7
Box A	58%	45%	36%	36%
Box B	10%	14%	13%	6%
No difference	32%	41%	51%	58%

The third option is the correct response for item 26.

The pattern of errors on this item is similar for all grades. The majority of children who made an error selected Box A. They probably selected this option because there are more cards with "X" on them in Box A than in Box B.

27. Spinner A and Spinner B are used to play this game. The spinners are marked as in the picture.

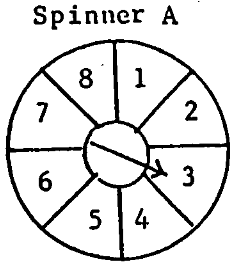
To play this game you spin the arrow on one of the spinners. (If the arrow stops on a line you spin it again.)

You win if the arrow points to a space with a "5" on it when it stops.

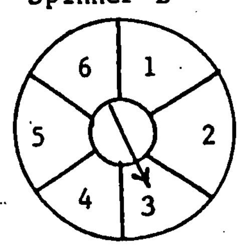
You lose if the arrow points to a space with any other number on it.

If you can play this game only once, which spinner would you choose to use so that you would have the better chance of winning?

Spinner A



Spinner B



_____ Spinner A

_____ Spinner B

_____ It doesn't make any difference

Table 109

Per Cent of Children Who Chose Each Option on Item 27

	Grade 4	Grade 5	Grade 6	Grade 7
Box A	23%	12%	7%	8%
Box B	38%	58%	70%	62%
No Difference	38%	30%	23%	30%

Spinner B is the correct response for item 27.

The results on this item are somewhat surprising. It was expected that this would be a very easy item. The third option was the most popular of the two incorrect options. The children may have picked this option because the spinners are the same size, disregarding the size of the individual spaces on the spinners.

28. Box A and Box B are used to play this game. The boxes contain black and white chips as in the picture.

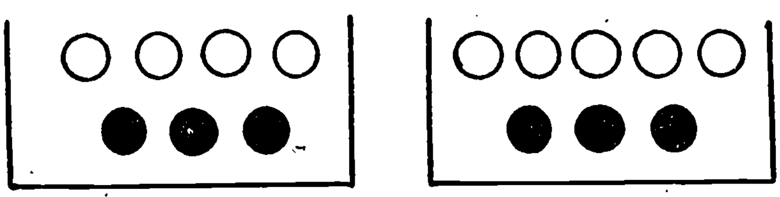
To play this game you pick a chip from one of the boxes without looking.

<u>You win if you pick a black chip</u>.

You lose if you pick a white chip.

If you can play this game only once, which box would you choose to pick from so that you would have the better chance of winning?

BOX A BOX B



_____ Box A

_____ Box B

_____ It doesn't make any difference

Table 110

Per Cent of Children Who Chose Each Option on Item 28

	Grade 4	Grade 5	Grade 6	Grade 7
Box A	48%	63%	73%	71%
Box B	22%	17%	9%	11%
No Difference	29%	20%	18%	18%

Box A is the correct response for item 28.

The patterns of responses for grades six and seven are almost identical. The item was more difficult for grades four and five but in each grade the third option was a better distractor than the second option. Children probably selected the third option because there are three black chips in each box disregarding the number of white chips in each box.

It is not clear why children would pick Box B as the better choice except perhaps as a random choice if they did not understand the item.

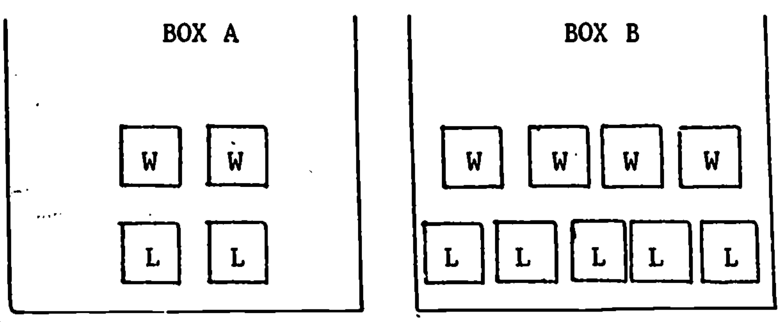
29. Box A and Box B are used to play this game. The boxes contain cards as in the picture.

To play this game you pick a card from one of the boxes without looking.

<u>You win if you pick a card with a "W" on it</u>.

You lose if you pick a card with a "L" on it.

If you can play this game only once, which box would you choose to pick from so that you would have the better chance of winning?



_____ Box A

_____ Box B

_____ It doesn't make any difference

Table 111

Per Cent of Children Who Chose Each Option on Item 29

	Grade 4	Grade 5	Grade 6	Grade 7
Box A	34%	56%	67%	67%
Box B	50%	31%	25%	25%
No difference	15%	12%	8%	8%

Box A is the correct response for item 29.

Box B was a much better distractor than the third option. In grade four, 50% of the children selected option B. It can be assumed that children picked Box B because it contains four winning cards while Box A contains only two winning cards. The results of this item, as in the previous items, suggest that children are basing there answers on the number of winning counters in the boxes rather than on the probability of selecting a winning card, chip, ball, etc.

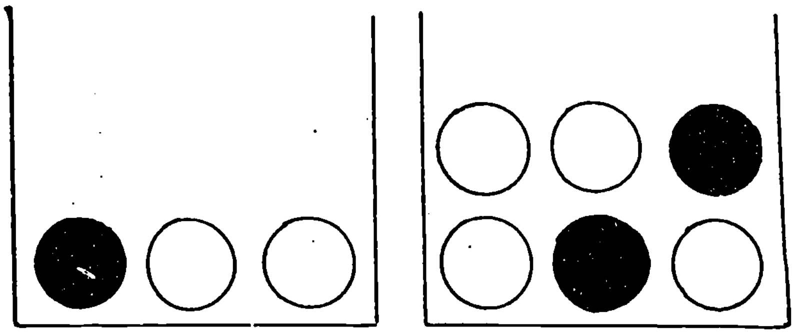
30. Box A and Box B are used to play this game. The boxes contain black and white chips as in the picture.

To play this game you pick a chip from one of the boxes without looking.

<u>You win if you pick a white chip</u>.

You lose if you pick a black chip.

If you can play this game only once, which box would you choose to pick from so that you would have the better chance of winning?



____ Box A

____ Box B

____ It doesn't make any difference

Table 112

Per Cent of Children Who Chose Each Option on Item 30

	Grade 4	Grade 5	Grade 6	Grade 7
Box A	24%	35%	37%	27%
Box B	62%	50%	45%	42%
No difference	13%	15%	16%	31%

The third option is the correct response for item 30.

This item was very difficult for all grades. Box B was the better distractor of the two incorrect responses. This is consistent with the pattern of errors on the other items in Test III.

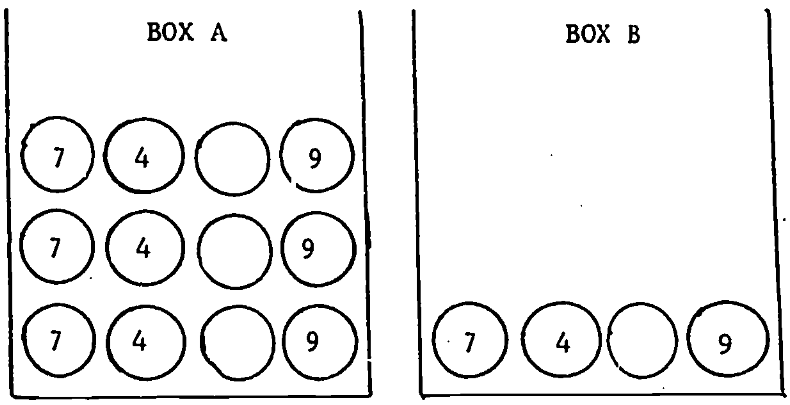
31. Box A and Box B are used to play this game. The boxes contain chips as in the picture.

To play this game you pick a chip from one of the boxes without looking.

<u>You win if you pick a blank chip</u>.

You lose if you pick a chip with a number on it.

If you can play this game only once, which box would you choose to pick from so that you would have the better chance of winning?



_____ Box A

_____ Box B

_____ It doesn't make any difference

Table 113

Per Cent of Children Who Chose Each Option on Item 31

	Grade 4	Grade 5	Grade 6	Grade 7
Box A	48%	44%	23%	24%
Box B	27%	26%	33%	18%
No difference	25%	29%	42%	58%

The third option is the correct response for item 31.

This item is very similar to item 30 but one would suspect that it would be more difficult than item 30. This is not true. Perhaps the orderly arrangement of the objects in the boxes may have made this item slightly easier.

As in the other items on this test, the box containing the greater number of winning counters was the better distractor.

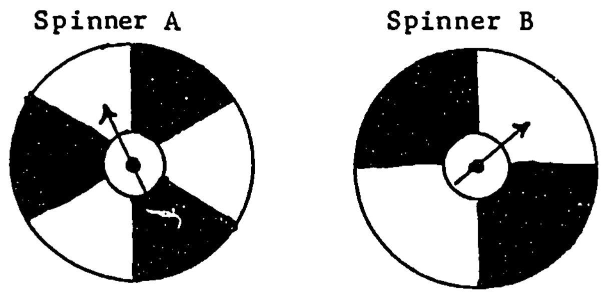
32. Spinner A and Spinner B are used to play this game. The spinners are marked as in the picture.

To play this game you spin the arrow on one of the spinners.

You win if the arrow points to a black part of the spinner when it stops.

You lose if the arrow points to a white part of the spinner.

If you can play this game only once, which spinner would you choose to use so that you would have the better chance of winning?



_____ Spinner A

_____ Spinner B

_____ It doesn't make any difference

Table 114

Per Cent of Children Who Chose Each Option on Item 32

	Grade 4	Grade 5	Grade 6	Grade 7
Box A	52%	33%	29%	27%
Box B	18%	32%	28%	17%
No difference	30%	35%	42%	56%

The third option is the correct response for item 32.

Over half of the children in the fourth grade and many children in the other grades selected Spinner A. Children probably made this choice because Spinner A contains three shaded regions while Spinner B has only two.

One can not be certain why children selected the third option. They may have recognized that the shaded areas on the two spinners are congruent or they may have made this choice because the spinners are the same size.

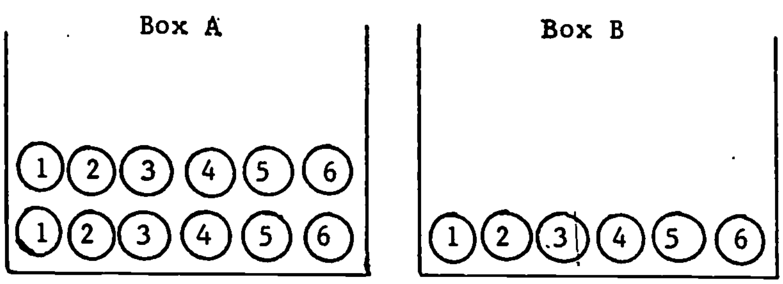
33. Box A and Box B are used to play this game. The boxes contain chips as in the picture.

To play this game you pick a chip from one of the boxes without looking.

<u>You win if you pick a chip with a "3" on it</u>.

You lose if you pick a chip with any other number on it.

If you can play this game only once, which box would you choose to pick from so that you would have the better chance of winning?



_____ Box A

_____ Box B

_____ It doesn't make any difference

Table 115

Per Cent of Children Who Chose Each Option on Item 33

	Grade 4	Grade 5	Grade 6	Grade 7
Box A	60%	40%	36%	34%
Box B	11%	18%	11%	9%
No difference	28%	41%	51%	55%

The third option is the correct response for item 33.

This item is similar to item 31 and functioned in much the same way except that it was somewhat easier for grades five and six.

Box A, with the greater number of winning chips, was the better distractor of the two incorrect options. This is the same type of error that is evident in the other items on Test III.

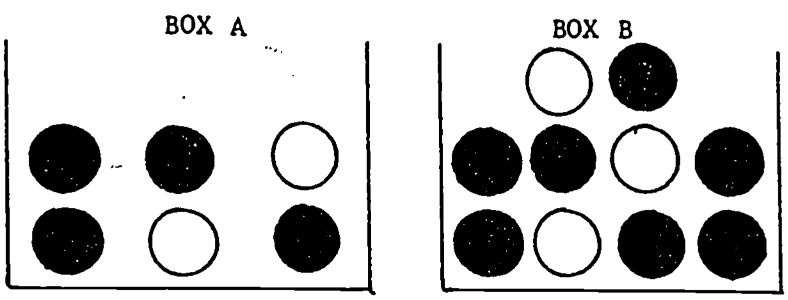
34. Box A and Box B are used to play this game. The boxes contain black and white chips as in the picture.

To play this game you pick a chip from one of the boxes wihtout looking.

<u>You win if you pick a white chip</u>.

You lose if you pick a black chip.

If you can play this game only once, which box would you choose to pick from so that you would have the better chance of winning?



_____ Box A

_____ Box B

_____ It doesn't make any difference

Table 116

Per Cent of Children Who Chose Each Option on Item 34

	Grade 4	Grade 5	Grade 6	Grade 7
Box A	34%	45%	48%	48%
Box B	48%	37%	31%	28%
No difference	18%	17%	19%	23%

Box A is the correct response for item 34.

The results of this item are surprising for it was assumed, based on the results of the pilot studies, that it would be the most difficult item on Test III. It can not be considered an easy item but the results in Table 116 show that it was considerably easier than item 30 (see Table 112). Items 34 and 30 are similar for both have figures in which the black and white chips are not arranged in an orderly manner. It is not clear how children made their decisions on this item.

Summary of Errors on Test III

From the responses on Test III there seems to be one main source of error on these items. It appears that many children made their decisions on the basis of the number of winning objects in the boxes rather than relating the number of winners in each box to the total number of objects in the box.

Chapter VI

SUMMARY AND CONCLUSIONS

The purpose of this study is to examine the status of three basic concepts of probability possessed by children in the fourth, fifth, sixth and seventh grades. The three concepts under investigation were: points of a finite sample space; probability of a simple event in a finite sample space; and quantification of probabilities.

The study was carried out during the first semester of the 1967-1968 academic year in the Wausau, Wisconsin, Public School System. The population for the study consisted of all children enrolled in the fourth, fifth, sixth and seventh grades in the Wausau district for whom a Total I.Q. on the California Test of Mental Maturity was available from the central office files. The population was limited to those children who had not had any formal learning experiences with topics in probability. The population included approximately 87% of the total number of children enrolled in grades four through seven in the district in October, 1967. The sample for the study consisted of 528 children randomly selected from the population. The children in the sample were categorized into twenty-four subgroups on the basis of sex, three I.Q. ranges and four grade levels.

Three tests, one for each of the three concepts listed above, were constructed by the writer for use in the study. Each test consisted of a set of items for which the child's responses would indicate if he could apply the concept in a variety of simple experiment and game situations.

Test I consisted of twelve items on the concept of sample space. The first six items (Subtest I-A) involved only simple counting. The last six items (Subtest I-B) involved simple ideas of combinations.

Test II consisted of twelve items on the concept of probability of a simple event. Each item on Test II presented a lot-drawing situation very similar to the situation presented in the corresponding item on Test I. The first six items on Test II (Subtest II-A) tested the notion of probability of a simple event in which the underlying ideas of sample space involved only simple counting. The last six items (Subtest II-B) tested the notion of probability of a simple event in which the underlying ideas of sample space involved combinations.

Test III consisted of ten items on the concept of quantification of probabilities. Each item presented a game situation in which the child had to decide which of two conditions represented the better probability of success for a specified simple event in one trial. Five of the items presented situations in which the specified event had the same probability of success under both conditions.

The tests were administered as written tests to groups of subjects during November and December, 1967. The same tests were administered to all subjects, grades four through seven. The items on all tests were scored either right or wrong.

A multivariate analysis of covariance was performed on the results of the three tests to determine whether significant differences occurred among the performances of children in the three I.Q. groups, two sex groups and four grade levels. Grade equivalent scores on the three parts of the Stanford Arithmetic Achievement Test were used as covariates. A univariate analysis of covariance was also performed on each of the dependent variables to determine the level of internal differences for significant overall effects. A discriminant function was calculated for each factor on which the F-statistic indicated significant variation among the mean vectors for the factor.

Correlation coefficients were calculated to determine which of three available scores obtained on the California Test of Mental Maturity, (Language I.Q., Non-Language I.Q. and Total I.Q.) was the best predictor of performance on three probability tests. Correlation coefficients were also calculated to gain some insight into what relationships exist between the children's performances on the three tests and subtests. An analysis of the errors that children made on each of the test items was performed in an attempt to determine what misconceptions the children may have about the concepts tested.

Results of the Study

In Chapter V the analysis of the data is presented in three parts: 1) the testing of hypotheses; 2) correlation studies; and 3) analysis of incorrect responses on the test items. A summary of the results of the study, obtained from these analyses, will be presented in a similar manner.

Part 1: Results of the Tests of Hypotheses

In a multivariate sense the overall mean performances, adjusted for the covariates, were significantly different ($p < .01$) among I.Q. groups, sex groups and grade levels. There were no significant interactions.

A further analysis of the differences among I.Q. groups revealed that the significant variation among mean vectors could be attributed to the significant differences among means on all three of the probability tests. The mean performances on all three tests ranged from high for the high I.Q. group to low for the low I.Q. group.

The variations among the mean vectors for boys and girls can be attributed to the marginally significant mean differences on Test I and Test III. On both of these tests the adjusted mean scores for girls are higher than the mean scores for the boys. The boys' performance on Test II was slightly better than the girls', but the difference is not significant.

The univariate F statistics for the four grade levels revealed that the significant variation among the mean vectors for grades was due mainly to the significant mean differences on Test I. The mean differences for grades on Tests II and III are not significant. While not all of the mean differences on the three tests are significant for the four grades, the direction of these differences is constant. The adjusted mean performances of the children in the four grades ranged from high for the seventh grade to low for the fourth grade over all tests.

The discriminant function computed for each of the significant main effects presented another way of characterizing the multivariate differences among groups.

The discrimination between I.Q. groups is an overall effect with Test I contributing most to the function. The function discriminates best between the high and low groups. It seems plausible to say that the high I.Q. group displays a better understanding of all three probability concepts than the low group. In particular, the high group has a better grasp of the notion of sample space.

The discriminant function for the main effect of sex indicates that the source of the difference between boys and girls is a contrast of Test I with Test III, sample space vs. quantification of probabilities. However, the distributions of discriminant scores for boys and girls are almost identical, indicating that the discriminant function for sex does not differentiate very well between boys and girls. Therefore, although the difference between the overall performances for boys and girls was statistically significant, the actual difference between groups is very small.

The discrimination among grades is due primarily to Test I, sample space. The function discriminates best between grade four and grades six and seven. Children in grades six and seven demonstrate a better understanding of sample space, particularly the associated notions of combinations, than children in grade four.

Part 2: Results of the Correlation Studies

The correlations of Language I.Q. and Non-Language I.Q. with the scores on the probability tests are very similar for all grades. The differences are all very small except for Test II in grade six. For this test the correlation with Language I.Q. is .48 and with Non-Language I.Q. the correlation is .16. The difference between these correlations is significant ($p < .01$).

In grade four Total I.Q. has the smallest correlations with the probability test scores. For grades five through seven the correlations of the test scores with Total I.Q. are slightly higher than the correlations with either Language I.Q. or Non-Language I.Q. except for Test II in grades six and seven. Although the differences are small this result give some support to the writer's assumption that total I.Q. would be the best predictor of performance on the probability tests. It was on the basis of this assumption that Total I.Q. was selected as one of the stratifying variables for the study.

The correlations between total performance scores on the three tests within grades are all significantly different from zero. The correlation coefficients range from .40 to .58 with the majority between .40 and .50. (See Tables 74-76.) This indicates that the tests are interdependent as would be expected. The concept of probability of a simple event certainly involves the concept of sample space. The concept of quantification of probability involves the idea of sample space and the idea of probability of a simple event. The results of these correlation studies do not indicate that these tests can be collapsed even though they are related. On the basis of the data from this study it appears that one can not be certain that a child understands the idea of sample space because he answers some questions about probability of simple events correctly. This is apparent from the results of correlation studies between Subtest I-A and II-A. Also, the results of Test III would not give a very clear picture of what the child understands about sample space and probability of a simple event.

On the basis of total mean scores, Test I (sample space) was the easiest for all grades and Test II (probability of a simple event) was most difficult for all grades.

Subtests I-A and II-A consist of the first six items on Tests I and II respectively. The mean scores on Subtest I-A are 3.56, 4.20, 4.77 and 5.00 and the mean scores on Subtest II-A are 2.28, 2.65, 3.55 and 3.97 for grades four through seven respectively. Subtest I-A was relatively easy for all grades. Subtest II-A was more difficult but still relatively easy for grades six and seven. The correlations between total scores on these subtests are .42, .25, .43 and .47 for grades four through seven respectively. Because of the close relationship between corresponding items on the subtests and the reasonably high mean scores, particularly for grades six and seven, these correlations are lower than expected.

An even more surprising result is the pattern of very low inter-item correlations between pairs of corresponding items on the subtests. Over 30% of these correlations are not significantly different from zero. Only the items involving sampling without replacement show a significant interdependent relationship for all grades. This relationship may be attributed to the difficulty in understanding the notion of sampling without replacement. (See Tables 77-79).

It seems reasonable to say that children in all grades exhibit considerable knowledge and understanding about sample space involving only simple counting. Some children in all grades also have an understanding of probability of a simple event and can apply this knowledge in a variety of different situations. However, the ability to answer a question about the probability of a simple event does not necessarily indicate that the child also recognizes all the elements of the sample space which contains the event.

Subtests I-B and II-B consist of the last six items on Test I and II respectively. The mean scores on Subtest I-B are 1.92, 2.97, 3.80 and 4.00; and the mean scores on Subtest II-B are .62, .66, .97 and 1.18 for grades four through seven respectively. The correlations between total scores on these subtests are .27, .32, .29 and .27 for grades four through seven respectively. The inter-item correlations for the six pairs of corresponding items on the subtests were all very low. Eighty per cent of these correlations are not significantly different from zero with the other 20% being significant at the .05 level. (See Tables 80-82).

The low correlations between total scores and low correlations between pairs of corresponding items on these subtests are undoubtedly due primarily to the extreme difficulty of Subtest II-B.

These results indicate that children in the fifth, sixth and seventh grades demonstrated some understanding of combinations in situations involving the idea of sample space but they were not able to apply these ideas in similar situations involving the probability of a simple event. Children in fourth grade had considerable difficulty with these items and apparently have not developed the ability to work with combinations to any great extent.

Part 3: Results of the Analyses of Incorrect Responses on the Test Items

The patterns of errors on each of the test items are discussed in Chapter V. The analyses of errors suggest several generalizations about misconceptions children may possess, and some apparent difficulties children experienced with interpreting the wording of some of the items.

There was some misunderstanding about the meaning of the word "different" as it was used in the context of items on Test I. Children were not able to decide when outcomes for an experiment were different and when they were the same. An expression like "pick a card from each box," in items involving combinations, was also misinterpreted by many children. Either the children ignored the word "each" or they did not understand what it implied. These children included pairs of objects from the same box in their answers as well as pairs of objects obtained by selecting one object from each box. The difficulties with these words was most apparent for children in the fourth grade although some children in all grades experienced these problems.

The idea of sampling without replacement was a difficult idea for many children at all levels. The common error was to still consider the object that was removed as a possible outcome for the experiment.

Many of the errors on the items in Test I which asked the child to list all possible combinations were due to the child omitting just one or two of the combinations. In general it would seem that these children did not have a systematic method for generating the combinations. This result is consistent with one of Piaget's findings. He concluded that children in the concrete operations stage (ages 8-11) demonstrated some understanding of combinations and permutations but had not yet developed a systematic method for computing such arrangements.

Many children do not distinguish between the probability of success for an event and the odds for success for the same event. This was very apparent in the answers for items on Subtest II-A. In general, children who gave the odds for winning rather than the probability of winning were not consistent in their use of this representation. This indicates that they probably were not aware of the difference between the two types of statements. Very few children used the odds representation for items which involved combinations. This was probably due in part to the fact that these items were extremely difficult. Also, for the majority of these items, the number of ways of winning was greater than or equal to the number of ways of losing. Therefore the odds for winning had to be expressed something like "3 out of 1" because of the specified form for the answer. It is assumed that children would tend to avoid giving this type of answer because it doesn't sound right.

The confusion between "odds" and "probability" is consistent with the findings of Leake in his study with junior high students.

Another source of error, that was apparent in items involving combinations, was that children often based their answers on the number of objects in the diagrams, or on the juxtaposition of the objects, rather than on the number of possible combinations.

On Test III (quantification of probabilities) the most common error appeared to be that children based their answers on the number of "winners" in the boxes rather than relating the number of winning objects in each box to the total number of objects in the box. Fourth and fifth grade children tended to make this error more often than children in the sixth and seventh grades. The items on Test III are very similar to the experiments used by Piaget in his study of quantification of probabilities. The main difference is that Piaget used concrete objects in an interview situation while these items were presented in written form. The results of this test agree with the conclusions of Piaget.

Conclusions and Implications

The results of this study can only be generalized for the population under consideration. The study would have to be replicated for other populations before the conclusions, which are based on the results of the study, can be considered to apply to other than the fourth, fifth, sixth and seventh grades of the Wausau, Wisconsin, Public School District.

Status of the Concepts

The most significant outcome of this study is that the children demonstrated that they had acquired considerable knowledge about the three concepts of probability under investigation and could apply these concepts in a variety of situations. These children had not received formal training on the notions of probability so their understanding and ability to apply these concepts must have developed as a result of their background, experiences and intuition.

This result supports the findings of Piaget, Leake and others, that young children do acquire some basic understanding about concepts of probability outside of school.

Implications

Several implications for educational practice and research are suggested by this result. These implications and research problems will be considered in the discussion that follows.

The most important implication for educational practice that arises from this study is that, since young children acquire some knowledge of probability outside of school, it seems reasonable to assume that some topics of probability would not be too difficult to include in the elementary school program.

If probability is to be included in the elementary school curriculum two questions that arise immediately are: 1) What topics of probability are most appropriate for the elementary school? 2) When should these topics be introduced? Both questions are very difficult to answer and perhaps have no definitive answers. Certainly more research is needed before one can attempt to answer these questions on the basis of experimental evidence. The results of this study give no conclusive answers to these questions but do suggest some topics that might be considered appropriate. The results also give some indication about the relative difficulty of the concepts included in the study for children at different levels. This provides some information that can serve as a guide for the placement of certain ideas of probability.

It appears that the concept of sample space can be included as early as grade four. The notion of sample space involving combinations was difficult for fourth grade children but this does not mean that it should not be included in the curriculum. The most serious difficulty seemed to be a lack of a system for generating the combinations. An interesting problem that is suggested by this error is: How early can children be taught to effectively use a systematic method for computing combinations? This question is one that should be studied because it has implications for use of materials on probability that are currently available to the schools. The SMSG materials on probability for the primary grades include activities involving simple combinations. The authors of these materials presume that children in the primary grades can be taught systematic methods for obtaining combinations. It is important to note that it is not necessary to force the notion of combinations for many interesting and worthwhile problems about probability can be posed without including combinations.

The fact that young children appear to know quite a bit about sample space involving simple counting supports the assumption underlying many of the activities and exercises on probability that are suggested for use in the elementary school. These activities generally do not include preliminary work with listing the elements of a sample space, assuming that children already have a good grasp of this notion. Even though this study supports this assumption, teachers who use such materials should not take this ability for granted. This study also shows that children do have some difficulty with counting and listing the points of a sample space. These children would undoubtedly profit by more practice with exercises in which they were specifically asked to list all of the points of a sample space.

The idea of probability of a simple event in which the underlying sample space involved only simple counting was relatively easy for all grades. Apparently this idea can also be successfully taught as early as grade four. Fourth grade children had more difficulty with the items than children in other grades but they did demonstrate some understanding of the basic idea of probability of a simple event.

There are several implications for teachers of this topic. Children must be taught the difference between "odds" and "probability." Also, one can not be certain that a child who gives a correct response to a question about the probability of a simple event actually recognizes all the elements of the sample space which contains the event. This latter point is one that is often assumed in the probability activities and exercises suggested for use in the elementary school. This study showed that this assumption is not necessarily true.

Probability of a simple event which involves combinations was an extremely difficult topic for all grades. Young children may be able to understand this idea after training but this needs further study.

The results of the test on quantification of probabilities agree with the findings of Piaget that children often base their answers on the number of "winners" rather than on the probabilities of success under different conditions. It is not clear from the children's written responses how they actually think about such items or how they decide which answer is correct. Further study with the type of item used in Test III, using an interview technique, could provide valuable information about children's understanding of this concept as well as a deeper insight into their understanding of sample space and probability of a simple event.

The Relationship of the Overall Performance on the Tests with the Factors of I.Q., Sex and Grade

All three of the main factors of this study; I.Q., sex and grade, are significantly related to the overall performance on the three tests.

The significant relationship between grade and overall performance gives substance to the proposal for introducing probability into the elementary school curriculum. Since understanding of ideas of probability seems to develop naturally in children as they grow older, it seems reasonable to suppose that this development could be increased and strengthened through formal learning activities during these early formative years.

The significant relationship between I.Q. and overall performance suggests that it may be necessary to differentiate the types of activities presented to different ability groups. This study does not indicate that low ability children can not or should not be taught topics of probability. It does indicate that low ability pupils have not acquired as much knowledge about the concepts investigated as the higher ability groups. This means that they will probably need more work with preliminary activities and their progress may be slower.

The relationship between sex and overall performance was significant but the internal causes for this relationship are not clear. This result does not suggest any implications for educational practice.

Suggestions for Further Study

The analysis of errors on the test items pointed out several misconceptions that may be artifacts of the tests or may represent deeper problems. These problems need further investigation before any conclusions can be drawn.

The possible misinterpretation of the words "different" and "each" should be tested in other settings. Are children confused because of the confounding of their use with probability notions or is there some confusion in the general use of these words?

It is not clear from the study how children interpreted questions involving the use of spinners. Does the size of the spinner make a difference in the way children view these items? This is an important question because many of the suggested materials for probability activities include the use of spinners. Does the shape of the spinner make a difference? It would be interesting to pose questions about two spinners in which one of the spinners was the usual disc and the other had a different shape such as an octagon, hexagon, etc.

A serious limitation of this study was that each test item was presented as a written item and the accompanying diagram could only depict the objects used in the problem in a static position. Would children give the same responses to these items if they had concrete objects to manipulate rather than having to use a picture to help interpret the problem? This problem could be investigated by randomly assigning subjects to one of two groups and using both techniques.

In order to gain a deeper insight into how children think about the items presented, an interview testing procedure could be used. In this way subjects could be asked to explain how they arrive at their answers. The information gained from this type of study would be very helpful in deciding how to present these topics to children in order to clarify misconceptions that they may already have formed.

It would also be valuable to expand the study of each concept by using a greater number of items and present a greater variety of situations. This should be done within grades and across grades.

This study should be replicated or similar studies conducted for other populations to see if consistent results are obtained.

Another very broad area for further study is: What effect will the teaching of these concepts have on the performance of children in applying these concepts? This involves the construction of appropriate units forstudy and the development of evaluation instruments.

This study has provided some insight into the status of probability concepts in young children. It has pointed out several areas that should be given careful consideration in classroom practice if the topic of probability is included in the elementary school curriculum. It has also suggested a number of problems that require further study.

APPENDIX A

PROBABILITY TESTS I, II, AND III,

DIRECTIONS FOR ADMINISTERING THE TESTS
AND
ANSWER KEY

Directions for Administering Test I

Do not open the booklets until I tell you.

Print the following information on the front page of your booklet:

Your name

The name of your school

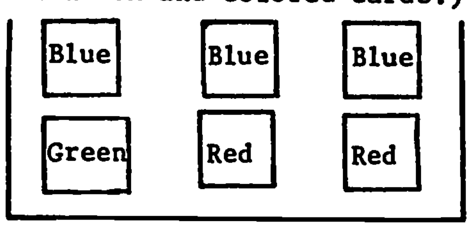
The grade you are in

In the first part of the booklet you will be asked to think about doing some experiments. These experiments will be like playing a game, where you do things like picking a card from a box without looking, spinning a spinner, or throwing a pair of dice. In each case you can not be certain what will happen each time you do the experiment. All of the things that can happen when you do an experiment are called the outcomes of the experiment. Different experiments will have different outcomes. For each question in this first part you will be asked to write the list of all of the outcomes that are possible for the experiment described in the question.

Look at the first sample question I have written on the chalkboard (card). (Read aloud and demonstrate with box and colored cards.)

A box contains cards as in the picture.

After I shake the box so that the cards are well-mixed, I pick one card from the box without looking.



The color of the card that I pick is called an outcome of this experiment.

(Do you understand that in this sample: The experiment is picking a card from the box without looking? An outcome of the experiment is the color of the card I pick?)

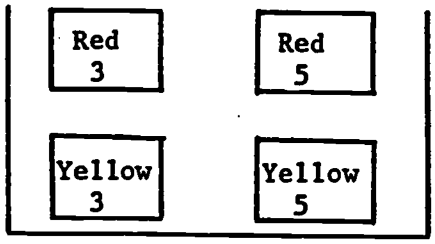
What are all the different outcomes that are possible for this experiment? (Write the outcomes on chalkboard or card: red, blue, green.)

Are there any questions about this first sample question?

Look at the second sample question I have written on the chalkboard (card). (Read aloud and demonstrate with box and colored cards.)

A box contains cards as in the picture.

After I shake the box so that the cards are well-mixed, I pick one card from the box without looking.



The color-number pair of the card that I pick is called an outcome of this experiment.

What are all the different outcomes that are possible for this experiment?

(Write outcomes on chalkboard or card: (red, 3), (red, 5), (yellow, 3), (yellow, 5))

Are there any questions about this second sample question?

Do you all understand what I mean by experiment and outcome of an experiment?

Now I am going to ask you to answer some questions like the sample questions. To begin with I will read each question aloud and I want you to read along with me, but you are to read silently.

Each question has a picture that goes with it. It is very important that you look at each picture carefully because you will need the picture to help answer the question. Think of the objects in the pictures as being well-mixed so that you would not be certain which object you would pick from the box if you pick without looking.

Write your list of outcomes for each question in the space under the question. You may do any scratch work that you may want to do on the test booklet or on your scratch paper.

Are there any questions before we begin?

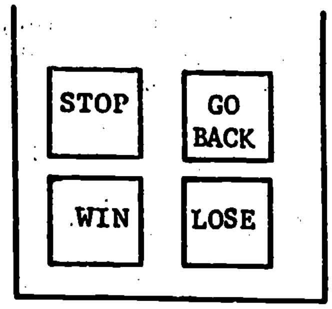
Open your booklets to page 1.

(Read questions 1 - 7 aloud. Allow enough time for all subjects to answer each question before reading the next question. Subjects are then to work at their own speed and stop after question 12.)

Stop when you get to the blue sheet in your booklet.

21. A box contains slips of paper as in the picture.

To play this game you pick three slips from the box at the same time without looking.



You win if one of the three slips that you pick has the word "WIN" on it.

You lose if you do not pick the slip with the word "WIN" on it.

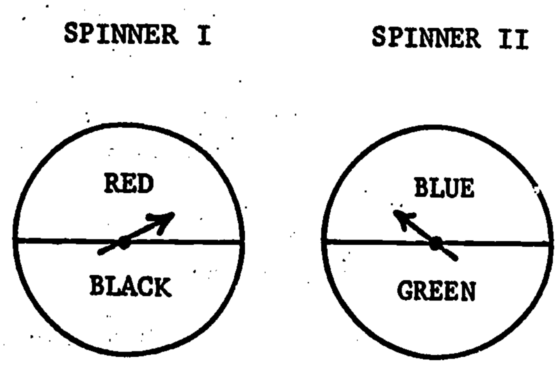
If you play this game only once, what chance do you have of winning?

Answer: _____ out of _____

****

22. Two spinners are marked as in the picture.

To play this game you spin the arrow on each of the spinners. (If an arrow stops on a line you spin it again.)



You win if the arrow on the first spinner points to a space marked red and the arrow on the second spinner points to a space marked blue when they stop.

You lose if the arrows stop in any other way.

If you play this game only once, what chance do you have of winning?

Answer: _____ out of _____

NAME ______________________________

GRADE ______________________________

SCHOOL ______________________________

DO NOT OPEN THIS BOOKLET

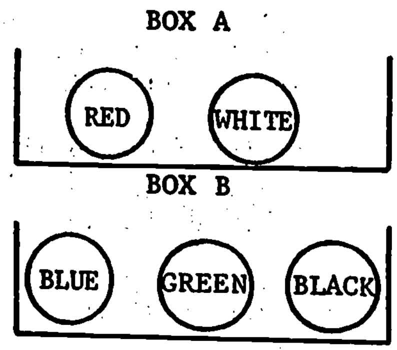
UNTIL YOU ARE TOLD TO DO SO

23. Box A and Box B contain colored chips as in the picture.

To play this game you pick one chip from each box without looking.

You win if you pick the red chip from Box A and the blue chip from Box B.

You lose if you do not pick this pair of chips.



If you play this game only once, what chance do you have of winning?

Answer: _____ out of _____

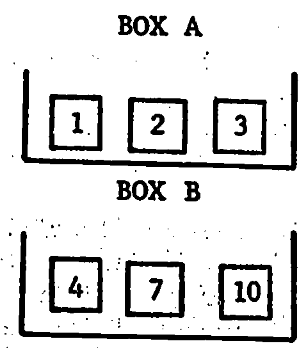
****

24. Box A and Box B contain cards as in the picture.

To play this game you pick one card from each box without looking.

You win if the sum of the numbers on the two cards is 6.

You lose if the sum is any other number.



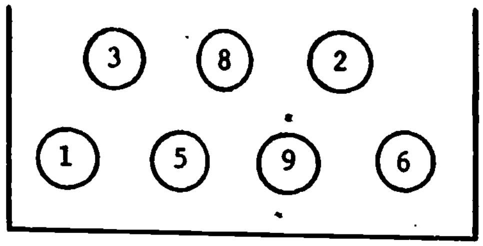
If you play this game only once, what chance do you have of winning?

Answer: _____ out of _____

1. For this experiment a box contains balls as in the picture.

To do this experiment you pick one ball from the box without looking.

The number that is on the ball that you pick is called an outcome of this experiment.



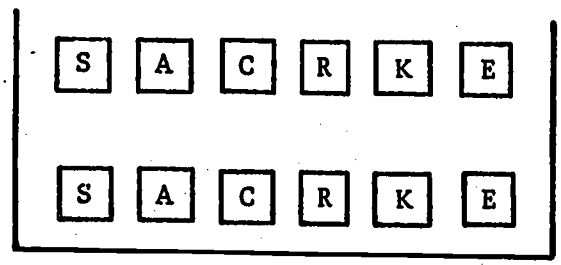
In the space below, write all the different outcomes it would be possible to obtain for this experiment.

****

2. For this experiment a box contains cards as in the picture.

To do this experiment you pick one card from the box without looking.

The letter that is on the card that you pick is called an outcome of this experiment.



In the space below, write all the different outcomes it would be possible to obtain for this experiment.

Directions for Administering Test III

In this third part you will be asked to think about some games. In each question you will be told how to play a certain game and how you can win the game.

Open your booklet to question 25, page 13.

(Read question 25 aloud while subjects read silently).

Remember, think of the objects in each box as being well-mixed. The picture only shows what is in each box, not the way in which the objects are placed in each box.

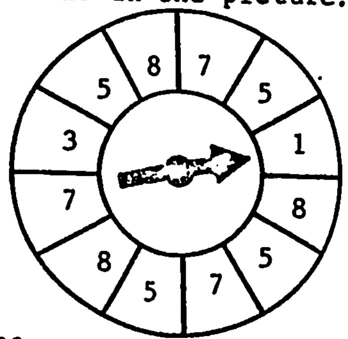
Answer this question by placing a check next to one of the three choices at the right of the picture. Remember you want to try to win the game. If you think you would have a better chance of winning by picking from Box A then put a check in the blank next to the words "Box A." If you think you would have a better chance of winning by picking from Box B then put a check in the blank next to the words "Box B." If you think it doesn't make any difference which box you pick from then put a check in the third blank, next to the words "it doesn't make any difference."

Are there any questions?

Now work ahead at your own speed and answer questions 26-34 in the same way you answered question 25. Check only one of the three blanks in each question.

When you have answered all of the questions you may return to your classroom. Leave your booklet and scratch paper on the table.

3. For this experiment a spinner is marked as in the picture.



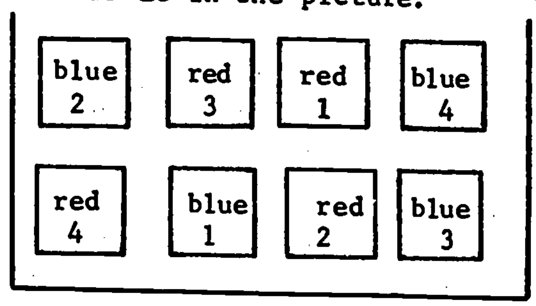
To do this experiment you spin the arrow on the spinner. (If the arrow stops on a line you spin it again.)

The number that the arrow points to when it stops is called an <u>outcome</u> of this experiment.

In the space below, write <u>all the different outcomes</u> it would be <u>possible to obtain</u> for this experiment.

****

4. For this experiment a box contains cards as in the picture.



To do this experiment you pick one card from the box without looking.

The <u>color-number pair</u> that is on the card that you pick is called an <u>outcome</u> of this experiment. For example, one outcome is the color-number pair (red, 3).

In the space below, write <u>all the different outcomes</u> it would be <u>possible to obtain</u> for this experiment.

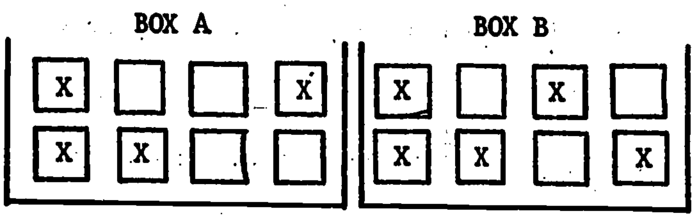
25. Box A and Box B are used to play this game.

The boxes contain cards as in the picture.

To play this game you pick a card from one of the boxes without looking.

<u>You win if you pick a card with a "X" on it.</u>
You lose if you pick a blank card.

If you can play this game only once, which box would you choose to pick from so that you would have the better chance of winning?



_____ Box A

_____ Box B

_____ It doesn't make any difference

****

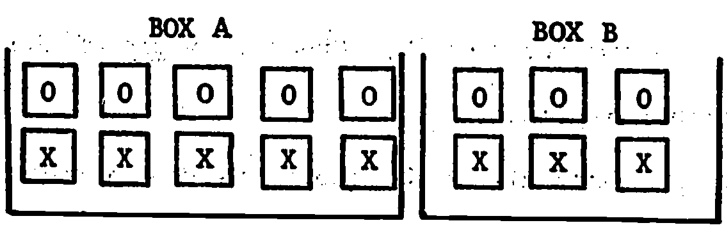
26. Box A and Box B are used to play this game.

The boxes contain cards as in the picture.

To play this game you pick a card from one of the boxes without looking.

<u>You win if you pick a card with an "O" on it.</u>
You lose if you pick a card with an "X" on it.

If you can play this game only once, which box would you choose to pick from so that you would have the better chance of winning?

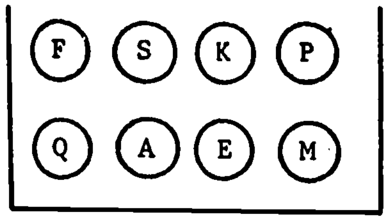


_____ Box A

_____ Box B

_____ It doesn't make any difference

5. For this experiment a box contains chips as in the picture.



To do this experiment you pick one chip from the box without looking.

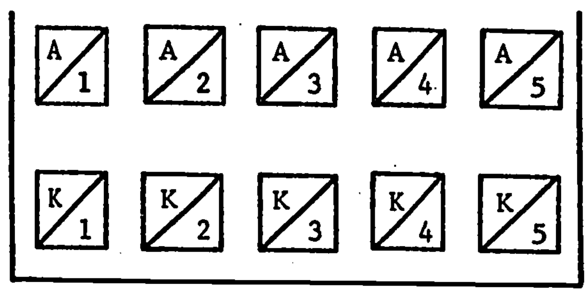
The letter that is on the chip that you pick is called an <u>outcome</u> of this experiment.

Imagine that the first chip you pick has a "K" on it. You <u>do</u> <u>not</u> put this chip back into the box. Then you pick a second chip.

In the space below, write <u>all</u> <u>the</u> <u>different</u> <u>outcomes</u> it would be possible to obtain for this experiment on the <u>second</u> try.

****

6. For this experiment a box contains cards as in the picture.



To do this experiment you pick one card from the box without looking.

The <u>letter/number</u> <u>pair</u> that is on the card that you pick is called an <u>outcome</u> of this experiment. For example, one outcome is the letter/number pair (K/4).

Imagine that the first <u>two</u> cards that you pick have the number "4" on them. You <u>do</u> <u>not</u> put these cards back into the box. Then you pick a third card.

In the space below, write <u>all</u> <u>the</u> <u>different</u> <u>outcomes</u> (letter/number pairs) that it would be <u>possible</u> <u>to</u> <u>obtain</u> on the <u>third</u> pick.

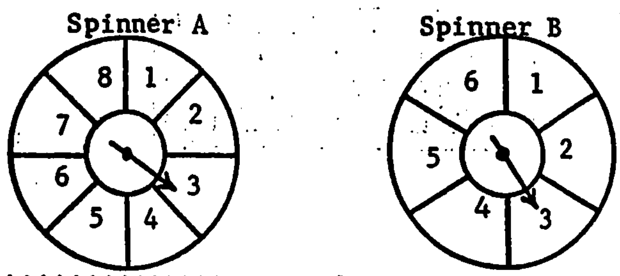
27. Spinner A and Spinner B are used to play this game.

The spinners are marked as in the picture.

To play this game you spin the arrow on one of the spinners. (If the arrow stops on a line you spin it again.)

You win if the arrow points to a space with a "5" on it when it stops.
You lose if the arrow points to a space with any other number on it.

If you can play this game only once, which spinner would you choose to use so that you would have the better chance of winning?



_____ Spinner A

_____ Spinner B

_____ It doesn't make any difference

****

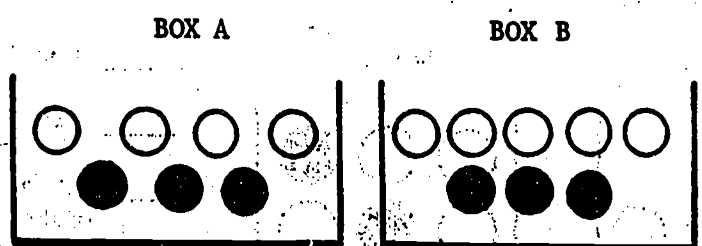
28. Box A and Box B are used to play this game.

The boxes contain black and white chips as in the picture.

To play this game you pick a chip from one of the boxes without looking.

You win if you pick a black chip.
You lose if you pick a white chip.

If you can play this game only once, which box would you choose to pick from so that you would have the better chance of winning?



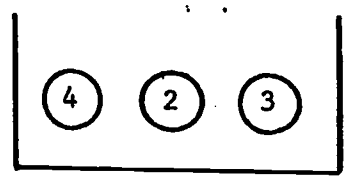
_____ Box A

_____ Box B

_____ It doesn't make any difference

7. For this experiment a box contains chips as in the picture.

To do this experiment you pick two chips from the box at the same time without looking.



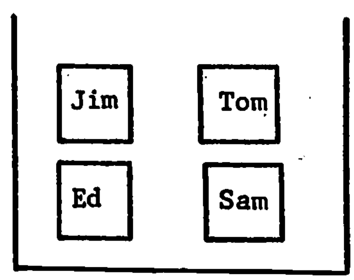
The sum of the numbers on the two chips that you pick is called an outcome of this experiment. For example, one outcome is the sum (4 + 2) or 6.

In the space below, write all the different outcomes (sums) it would be possible to obtain for this experiment.

****

8. For this experiment a box contains slips with names on them as in the picture.

To do this experiment you pick two slips from the box at the same time without looking.



The pair of names on the two slips that you pick is called an outcome of this experiment. For example, one outcome is the pair of names (Ed, Sam).

In the space below, write all the different outcomes (pairs of names) it would be possible to obtain for this experiment.

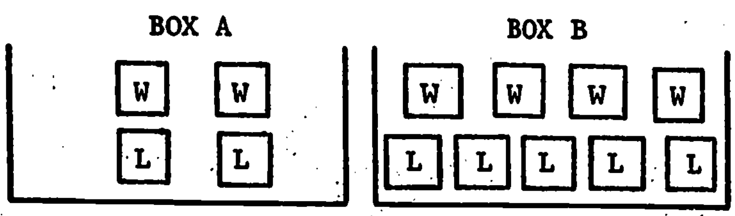
29. Box A and Box B are used to play this game.

The boxes contain cards as in the picture.

To play this game you pick a card from one of the boxes without looking.

<u>You win if you pick a card with a "W" on it.</u>
You lose if you pick a card with a "L" on it.

If you can play this game only once, which box would you choose to pick from so that you would have the better chance of winning?



_____ Box A

_____ Box B

_____ It doesn't make any difference

****

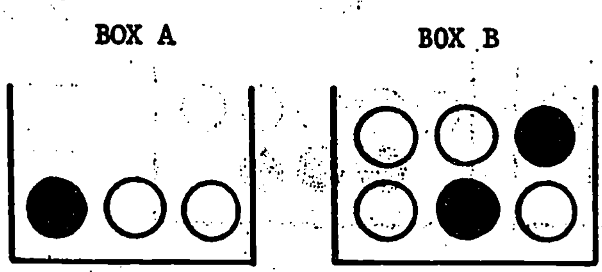
30. Box A and Box B are used to play this game.

The boxes contain black and white chips as in the picture.

To play this game you pick a chip from one of the boxes without looking.

<u>You win if you pick a white chip.</u>
You lose if you pick a black chip.

If you can play this game only once, which box would you choose to pick from so that you would have the better chance of winning?



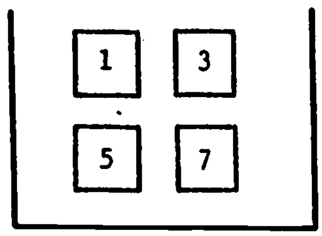
_____ Box A

_____ Box B

_____ It doesn't make any difference

9. For this experiment a box contains cards as in the picture.

To do this experiment you pick <u>three</u> cards from the box at the same time without looking.



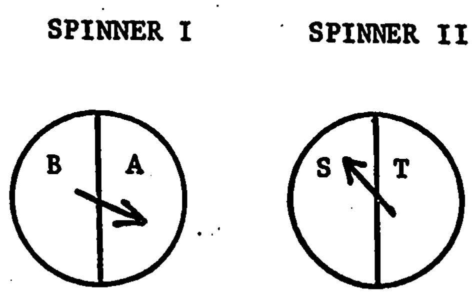
The sum of the numbers on the three cards that you pick is called an <u>outcome</u> of this experiment. For example, one outcome is the sum (1 + 3 + 7) or 11.

In the space below, write <u>all the different outcomes</u> (sums) it would be <u>possible to obtain</u> for this experiment.

****

10. For this experiment two spinners are marked as in the picture.

To do this experiment you spin the arrow on <u>each</u> spinner. (If an arrow stops on a line you spin it again.)



The <u>pair of letters</u> in the spaces that the two arrows point to when they stop is called an <u>outcome</u> of this experiment. For example, one outcome is the pair of letters (A, S).

In the space below, write <u>all the different outcomes</u> (pairs of letters) it would be <u>possible to obtain</u> for this experiment.

31. Box A and Box B are used to play this game.

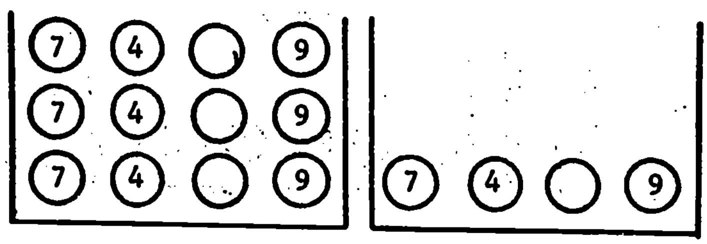
The boxes contain chips as in the picture.

To play this game you pick a chip from one of the boxes without looking.

<u>You win if you pick a blank chip.</u>
You lose if you pick a chip with a number on it.

If you can play this game only once, which box would you choose to pick from so that you would have the better chance of winning?

BOX A BOX B



_____ Box A

_____ Box B

_____ It doesn't make any difference

****

32. Spinner A and Spinner B are used to play this game.

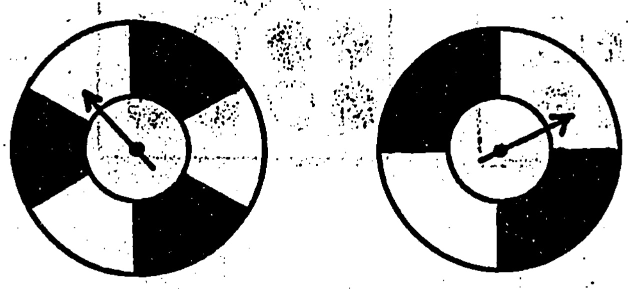
The spinners are marked as in the picture.

To play this game you spin the arrow on one of the spinners.

<u>You win if the arrow points to a black part of the spinner when it stops.</u>
You lose if the arrow points to a white part of the spinner.

If you can play this game only once, which spinner would you choose to use so that you would have the better chance of winning?

Spinner A Spinner B

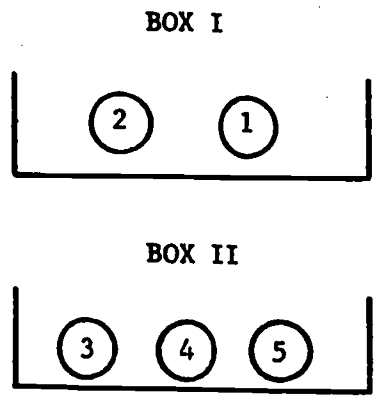


_____ Spinner A

_____ Spinner B

_____ It doesn't make any difference

11. For this experiment two boxes contain balls as in the picture.



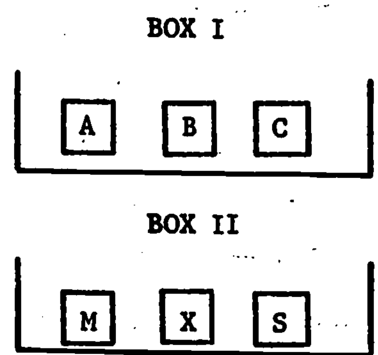
To do this experiment you pick one ball from each box without looking.

The product of the numbers on the two balls that you pick is called an outcome of this experiment. For example, one outcome is the product (2 x 4) or 8.

In the space below, write all the different outcomes (products) it would be possible to obtain for this experiment.

****

12. For this experiment two boxes contain cards as in the picture.



To do this experiment you pick one card from each box without looking.

The pair of letters on the two cards that you pick is called an outcome of this experiment. For example, one outcome is the pair of letters (C, X).

In the space below, write all the different outcomes (pairs of letters) it would be possible to obtain for this experiment.

33. Box A and Box B are used to play this game.

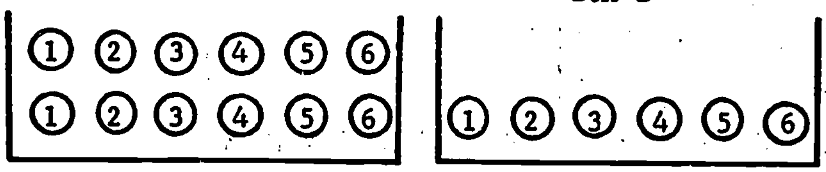
The boxes contain chips as in the picture.

To play this game you pick a chip from one of the boxes without looking.

You win if you pick a chip with a "3" on it.
You lose if you pick a chip with any other number on it.

If you can play this game only once, which box would you choose to pick from so that you would have the better chance of winning?

BOX A BOX B



_____ Box A

_____ Box B

_____ It doesn't make any difference

****

34. Box A and Box B are used to play this game.

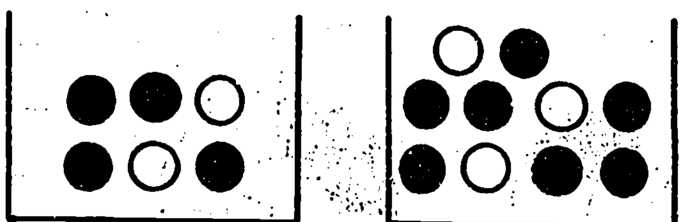
The boxes contain black and white chips as in the picture.

To play this game you pick a chip from one of the boxes without looking.

You win if you pick a white chip.
You lose if you pick a black chip.

If you can play this game only once, which box would you choose to pick from so that you would have the better chance of winning?

BOX A BOX B



_____ Box A

_____ Box B

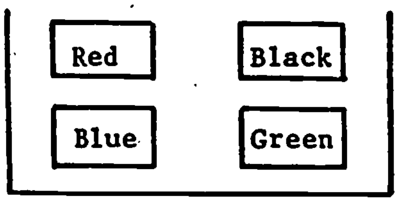
_____ It doesn't make any difference

Directions for Administering Test II

In this second part you will be asked to think about some games. In each question you will be told how to play a certain kind of game and how you can win that game. The question you will be asked is: what chance would you have of winning the game if you played the game only once?

Look at the sample question I have written on the chalkboard (card). (Read aloud and demonstrate with box and colored cards.)

A box contains four colored cards as in the picture. You win if you pick a red card. You lose if you pick any other card.



Answer: _____ out of _____

As you can see in the picture the box contains four colored cards; one red card, one blue card, one black card, and one green card. To play this game you shake the box so that the cards are well-mixed and then you pick one card from the box without looking. If you play this game only once, what chance do you have of winning the game?

You answer this question by filling in the blanks in the expression _____ out of _____.

(Ask subjects for answer. Write correct answer in blanks and discuss why answer must be 1 out of 4 and not 1 out of 3.)

Do you have any questions about the sample question?

Do you understand what is meant by chance of winning?

Do you understand how you are to answer this kind of question by filling in the blanks in the expression _____ out of _____?

Answer Key
Test I

1. 3, 8, 2, 1, 5, 9, 6

2. S, A, C, R, K, E

3. 1, 3, 5, 7, 8

4. (red,1), (red,2), (red,3), (red,4), (blue,1), (blue,2), (blue,3), (blue,4)

5. F, S, P, Q, A, E, M

6. (A/1), (A/2), (A/3), (A/5), (K/1), (K/2), (K/3), (K/5)

7. (4 + 2), (4 + 3), (3 + 2)

8. (Jim,Tom), (Jim,Ed), (Jim,Sam), (Tom,Ed), (Tom,Sam), (Ed,Sam)

9. (1 + 3 + 7), (1 + 5 + 7), (3 + 5 + 7)

10. (A,S), (A,T), (B,S), (B,T)

11. (2 x 3), (2 x 4), (2 x 5), (1 x 3), (1 x 4), (1 x 5)

12. (A,M), (A,X), (A,S), (B,M), (B,X), (B,S), (C,M), (C,X), (C,S)

TEST II

13. 1 out of 7

14. 2 out of 12

15. 3 out of 12

16. 1 out of 8

17. 1 out of 7

18. 5 out of 8

19. 2 out of 3

20. 1 out of 6

21. 3 out of 4

22. 1 out of 4

23. 1 out of 6

24. 1 out of 9

TEST III

25. Box B

26. It doesn't make any difference

27. Box B

28. Box A

29. Box A

30. It doesn't make any difference

31. It doesn't make any difference

32. It doesn't make any difference

33. It doesn't make any difference

34. Box A

Open your booklet to page 7, question 13. I will read this question aloud and you read along with me. You read silently while I read the question aloud. Remember, think of the objects in the picture as being well-mixed so that you would not be certain which object you would pick from the box if you pick without looking.

Now you work ahead at your own speed and stop after question 24. Stop when you get to the yellow sheet in your booklet.

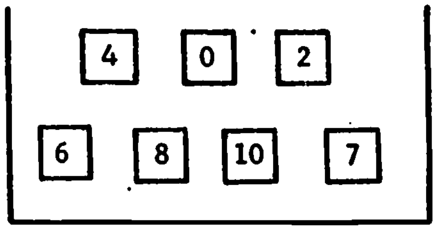
APPENDIX B

INTER-ITEM CORRELATIONS FOR SUBTESTS I-A AND II-A

FOR GRADES FOUR, FIVE, SIX AND SEVEN

13. A box contains cards as in the picture.

To play this game you pick one card from the box without looking.



You win if you pick the card with the "2" on it.

You lose if you pick a card with any other number on it.

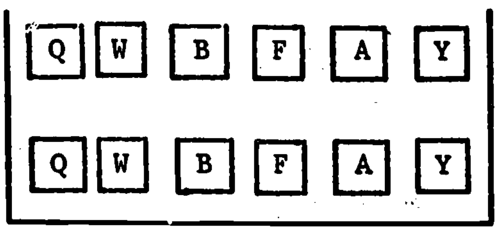
If you play this game only once, what chance do you have of winning?

Answer: _____ out of _____

****

14. A box contains slips of paper as in the picture.

To play this game you pick one slip of paper from the box without looking.



You win if you pick a slip with an "A" on it.

You lose if you pick a slip with any other letter on it.

If you play this game only once, what chance do you have of winning?

Answer: _____ out of _____

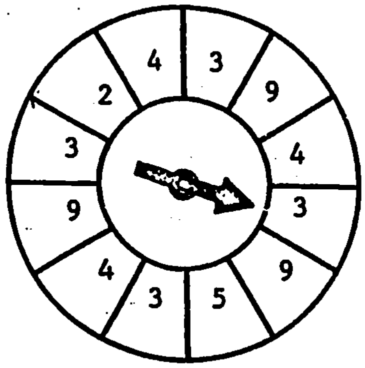
INTER-ITEM CORRELATION MATRIX OF SUBTEST I-A AND SUBTEST II-A FOR GRADE 4

	Subtest I-A						Subtest II-A					
	1	2	3	4	5	6	13	14	15	16	17	18
1	1.00	.18*	-.01	.27**	.27**	.27**	.13	.23**	.14	.19*	.18*	.17*
2		1.00	.54**	.01	.34**	.33**	.16	.19*	.11	.16	.15	.09
3			1.00	.02	.34**	.30**	.17*	.23**	.22**	.22**	.07	.12
4				1.00	.21*	.29**	.16	.24**	.21*	.10	.28**	.24**
5					1.00	.67**	.11	.21*	.13	.16	.29**	.23**
6						1.00	.30**	.37**	.30**	.20*	.37**	.39**
13							1.00	.29**	.25**	.57**	.21*	.22**
14								1.00	.58**	.14	.35**	.49**
15									1.00	.02	.42**	.57**
16										1.00	.09	.10
17											1.00	.53**
18												1.00

* $p < .05$ ** $p < .01$

15. A spinner is marked as in the picture.

To play this game you spin the arrow on the spinner. (If the arrow stops on a line you spin it again.)



You win if the arrow points to a space marked with a "4" when it stops.

You lose if the arrow points to a space with any other number on it.

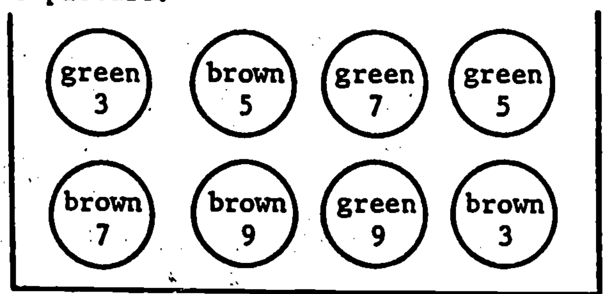
If you play this game only once, what chance do you have of winning?

Answer: _____ out of _____

****

16. A box contains balls as in the picture.

To play this game you pick one ball from the box without looking.



You win if you pick a green ball with a "5" on it.

You lose if you pick any other ball.

If you play this game only once, what chance do you have of winning?

Answer: _____ out of _____

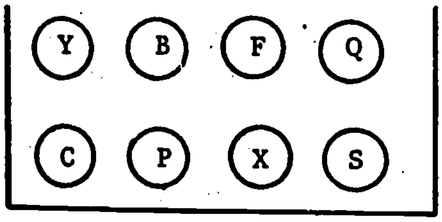
INTER-ITEM CORRELATION MATRIX OF SUBTEST I-A AND SUBTEST II-A FOR GRADE 5

	Subtest I-A						Subtest II-A					
	1	2	3	4	5	6	13	14	15	16	17	18
1	1.00	.19*	.31**	.23**	.02	.18*	.07	.18*	.14	.06	.21*	.17*
2		1.00	.54**	.09	.10	.08	.07	.05	.08	.10	.00	-.05
3			1.00	.04	-.02	.03	.12	.15	.08	.12	.04	.01
4				1.00	.23**	.32**	-.01	.20*	.29**	.12	.12	.23**
5					1.00	.65**	.08	.11	.12	.14	.19*	.15
6						1.00	.17*	.25**	.27**	.17*	.29**	.29**
13							1.00	.36**	.25**	.43**	.14	.26**
14								1.00	.58**	.17*	.29**	.44**
15									1.00	.09	.22**	.44**
16										1.00	.17*	.21*
17											1.00	.39**
18												1.00

* $p < .05$ ** $p < .01$

17. A box contains chips as in the picture.

To play this game you pick one chip from the box without looking.



You win if you pick the chip with the "X" on it.

You lose if you pick a chip with any other letter on it.

Imagine that the first chip you pick has "B" on it and is not a winner. You do not put this chip back into the box. Then you pick a second chip.

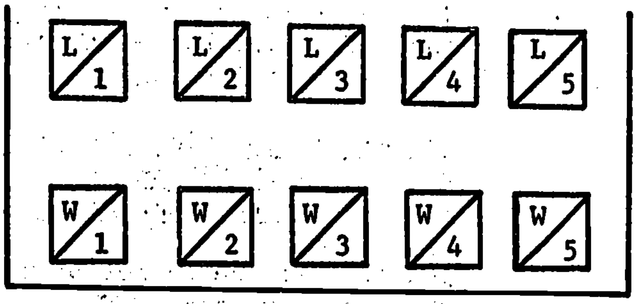
What chance do you have of winning on the second try?

Answer: _____ out of _____

****

18. A box contains cards as in the picture.

To play this game you pick one card from the box without looking.



You win if you pick a card with a "W" on it.

You lose if you pick a card with a "L" on it.

Imagine that the first two cards that you pick have "L" on them and are not winners. You do not put these cards back into the box. Then you pick a third card.

What chance do you have of winning on the third try?

Answer: _____ out of _____

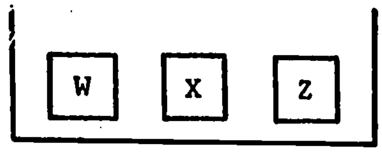
INTER-ITEM CORRELATION MATRIX OF SUBTEST I-A AND SUBTEST II-A FOR GRADE 6

	Subtest I-A						Subtest II-A					
	1	2	3	4	5	6	13	14	15	16	17	18
1	1.00	.12	.06	-.01	.04	.04	.19*	.07	-.01	-.01	.01	-.04
2		1.00	.66**	.05	.22**	.19*	.14	.13	.16	.18*	.17*	.11
3			1.00	.06	.19*	.21*	.15	.21*	.14	.11	.13	.17*
4				1.00	.15	.28**	-.06	.22**	.13	.25**	-.06	.07
5					1.00	.66**	.12	.27**	.21*	.10	.30**	.34**
6						1.00	.12	.25**	.20*	.21*	.22**	.36**
13							1.00	.25**	.10	.07	.15	.16
14								1.00	.62**	.09	.32**	.42**
15									1.00	.06	.30**	.47**
16										1.00	.15	.21*
17											1.00	.45**
18												1.00

* p < .05   ** p < .01

19. A box contains cards as in the picture.

To play this game you pick <u>two</u> cards from the box at the same time without looking.



<u>You win if one of the two cards that you pick is the card with the "X" on it.</u>

You lose if you do not pick the card with the "X" on it.

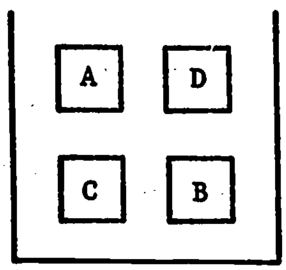
If you play this game only once, what chance do you have of winning?

Answer: _____ out of _____

****

20. A box contains cards as in the picture.

To play this game you pick <u>two</u> cards from the box at the same time without looking.



<u>You win if you pick the pair of cards with "A" on one card and "B" on the other card.</u>

You lose if you do not pick this pair of cards.

If you play this game only once, what chance do you have of winning?

Answer: _____ out of _____

INTER-ITEM CORRELATION MATRIX OF SUBTEST I-A AND SUBTEST II-A FOR GRADE 7

	Subtest I-A						Subtest II-A					
	1	2	3	4	5	6	13	14	15	16	17	18
1	1.00	.14	.06	.44**	.17*	.19*	.30**	.16	.07	.07	.22**	.16
2		1.00	.56**	.11	.26**	.14	.25**	.18*	.23**	-.06	.38**	.27**
3			1.00	.11	.35**	.26**	.10	.21*	.15	.17*	.46**	.34**
4				1.00	.09	.10	.23**	.24**	.18*	.20*	.14	.12
5					1.00	.63**	.07	.18*	.42**	-.07	.37**	.32**
6						1.00	.04	.24**	.37**	.09	.29**	.30**
13							1.00	.30**	.20*	.18*	.23**	.22**
14								1.00	.55**	.14	.32**	.39**
15									1.00	.02	.45**	.57**
16										1.00	.12	.04
17											1.00	.70**
18												1.00

* p < .05 ** p < .01

APPENDIX C

INTER-ITEM CORRELATIONS FOR SUBTESTS I-B AND II-B

FOR GRADES FOUR, FIVE, SIX AND SEVEN

INTER-ITEM CORRELATION MATRIX OF SUBTEST I-B AND SUBTEST II-B FOR GRADE 4

	Subtest I-B						Subtest II-B					
	7	8	9	10	11	12	19	20	21	22	23	24
7	1.00	.47**	.42**	.28**	.33**	.44**	.13	.07	-.01	.01	.08	.11
8		1.00	.48**	.25**	.34**	.40**	.22**	.11	.25**	.09	.10	.18*
9			1.00	.29**	.42**	.44**	.16	.06	.16	.11	.22**	.15
10				1.00	.40**	.49**	-.05	.11	.03	.00	.15	.19*
11					1.00	.61**	.08	.13	.05	.05	.16	.22**
12						1.00	.06	.12	.04	.02	.11	.20*
19							1.00	-.05	.43**	-.16	.05	-.08
20								1.00	-.04	.06	.30**	.66**
21									1.00	-.14	.02	.01
22										1.00	.35**	.14
23											1.00	.32**
24												1.00

* $p < .05$ ** $p < .01$

INTER-ITEM CORRELATION MATRIX OF SUBTEST I-B AND SUBTEST II-B FOR GRADE 5

	Subtest I-B						Subtest II-B					
	7	8	9	10	11	12	19	20	21	22	23	24
7	1.00	.24**	.44**	.21*	.26**	.26**	.19*	.08	.15	.05	.12	.12
8		1.00	.42**	.25**	.37**	.37**	.24**	.10	.16	.03	.16	.10
9			1.00	.35**	.40**	.36**	.22**	.12	.16	.12	.13	.22**
10				1.00	.55**	.59**	.15	.12	.05	.09	.16	.20*
11					1.00	.78**	.20*	.12	.18*	.08	.16	.11
12						1.00	.22**	.12	.19*	.03	.18*	.18*
19							1.00	.09	.55**	.16	.11	.14
20								1.00	.09	.19*	.31**	.46**
21									1.00	.07	.12	.05
22										1.00	.48**	.25**
23											1.00	.49**
24												1.00

INTER-ITEM CORRELATION MATRIX OF SUBTEST I-B AND SUBTEST II-B FOR GRADE 6

	Subtest I-B						Subtest II-B					
	7	8	9	10	11	12	19	20	21	22	23	24
7	1.00	.37**	.26**	.33**	.27**	.31**	.03	.02	.01	.15	-.06	.10
8		1.00	.31**	.20*	.25**	.28**	.11	.05	.07	.18*	.05	.10
9			1.00	.14	.19*	.16	.08	-.06	.04	.09	.01	.10
10				1.00	.66**	.74**	.13	.17*	.15	.04	.07	.14
11					1.00	.87**	.17*	.16	.20*	.16	.10	.08
12						1.00	.21*	.15	.16	.11	.09	.12
19							1.00	.09	.42**	-.09	.02	.05
20								1.00	.13	.01	.37**	.40**
21									1.00	-.10	-.01	.03
22										1.00	.47**	.30**
23											1.00	.49**
24												1.00

* $p < .05$ ** $p < .01$

INTER-ITEM CORRELATION MATRIX OF SUBTEST I-B AND SUBTEST II-B FOR GRADE 7

	Subtest I-B						Subtest II-B					
	7	8	9	10	11	12	19	20	21	22	23	24
7	1.00	.45**	.44**	.36**	.40**	.43**	.04	.11	.06	.03	.12	.11
8		1.00	.31**	.36**	.46**	.41**	.00	.16	.00	.11	.14	.14
9			1.00	.36**	.47**	.38**	.08	.16	.11	.13	.22**	.16
10				1.00	.63**	.61**	.09	.15	.08	.12	.16	.20*
11					1.00	.68**	.10	.09	.09	.02	.20*	.12
12						1.00	.03	.14	.01	.07	.24**	.18*
19							1.00	.15	.37**	.10	.08	.17*
20								1.00	.11	.28**	.56**	.72**
21									1.00	-.08	.03	-.02
22										1.00	.54**	.30**
23											1.00	.62**
24												1.00

* $p < .05$ ** $p < .01$

APPENDIX D

SCORES OF 528 CHILDREN ON THE CALIFORNIA TEST OF MENTAL MATURITY, STANFORD ARITHMETIC ACHIEVEMENT TEST AND THE PROBABILITY TESTS

California Test of Mental Maturity

L Language I.Q.

NL Non-language I.Q.

T Total I.Q.

Stanford Arithmetic Achievement Test (Grade Equivalent Scores)

S_1 Computation

S_2 Concepts

S_3 Applications

Probability Tests

I-A Sample Space (simple counting)

I-B Sample Space (combinations)

II-A Probability of a Simple Event (simple counting)

II-B Probability of a Simple Event (combinations)

III Quantification of Probabilities

Group ijk: I.Q. group i; Sex group j; Grade level k

I.Q. group 1: 71-104
I.Q. group 2: 105-113
I.Q. group 3: 114-144

Sex group 1: boys
Sex group 2: girls

Grade Level 1: grade four
Grade Level 2: grade five
Grade Level 3: grade six
Grade Level 4: grade seven

Group 111

Subject	I.Q. L	I.Q. NL	I.Q. T	Arith. Ach. S_1	Arith. Ach. S_2	Arith. Ach. S_3	Probability I-A	Probability I-B	Probability II-A	Probability II-B	Probability III
1	96	113	103	3.5	2.9	2.9	3	-	1	-	3
2	109	98	103	2.7	5.0	5.5	2	-	2	1	3
3	101	100	100	3.5	2.3	3.2	2	-	2	2	-
4	97	82	91	2.5	1.9	2.1	1	-	1	-	1
5	107	92	98	3.6	3.6	3.2	2	-	1	-	-
6	104	100	102	3.3	2.7	3.6	3	4	-	-	5
7	96	100	97	4.3	3.6	3.6	-	-	-	-	2
8	95	104	98	3.7	2.1	3.8	-	-	2	-	2
9	106	98	102	2.9	3.6	4.0	3	2	-	-	1
10	102	100	101	2.2	2.3	3.6	-	-	1	-	4
11	98	98	98	2.5	4.5	4.2	1	2	2	-	2
12	99	101	100	3.6	3.9	4.0	1	-	1	-	3
13	107	95	100	3.5	1.9	3.0	2	-	-	-	4
14	106	95	99	2.7	2.7	2.7	-	-	1	-	1
15	102	100	101	2.5	4.5	3.6	2	-	1	-	1
16	96	95	95	3.6	3.3	3.4	5	2	1	1	2
17	92	104	96	2.5	3.6	3.2	4	1	-	-	2
18	113	94	102	3.6	4.6	4.1	3	5	2	2	10
19	91	104	96	2.9	2.7	3.6	3	1	2	-	-
20	103	99	101	3.5	4.5	5.1	2	1	-	-	5
21	85	99	92	2.5	2.7	3.6	2	1	2	-	4
22	109	98	103	3.5	2.7	3.2	6	1	3	2	3

Group 112

Subject	I.Q. L	I.Q. NL	I.Q. T	Arith. Ach. S_1	Arith. Ach. S_2	Arith. Ach. S_3	Probability I-A	Probability I-B	Probability II-A	Probability II-B	Probability III
1	94	94	93	3.6	2.7	2.9	3	-	-	-	1
2	110	94	101	3.1	4.3	3.6	3	-	1	1	3
3	104	97	100	4.0	5.0	4.7	4	3	4	-	3
4	100	102	101	4.4	4.4	4.9	2	4	2	-	7
5	100	100	100	5.2	6.1	6.5	6	6	2	-	2
6	97	101	98	2.9	2.9	2.7	-	2	-	2	3
7	102	96	98	4.1	4.6	4.1	4	1	1	1	5
8	93	81	86	2.7	3.3	4.0	3	-	1	3	4
9	87	107	95	4.3	4.8	3.9	3	3	-	-	3
10	99	101	100	2.5	2.5	2.3	5	2	2	2	5
11	89	98	93	3.1	3.6	2.7	4	-	4	-	3
12	100	99	99	2.9	4.3	3.4	1	1	-	-	3
13	99	103	101	3.5	5.9	5.1	6	2	1	-	2
14	111	89	99	4.9	3.6	4.2	2	1	-	-	4
15	89	98	94	3.7	3.6	3.6	2	1	1	-	-
16	108	100	103	3.7	4.1	4.0	4	2	1	1	5
17	104	95	98	4.0	3.3	3.0	6	4	4	1	2
18	94	98	95	4.0	4.1	4.2	3	-	2	1	3
19	115	94	103	4.6	5.2	4.6	6	1	-	-	5
20	98	95	96	3.3	3.6	3.6	3	-	4	1	1
21	102	96	99	3.8	3.6	3.0	2	-	2	-	2
22	94	93	93	3.3	2.7	3.4	2	-	2	-	2

Group 113

Subject	I.Q. L	I.Q. NL	I.Q. T	Arith. Ach. S_1	Arith. Ach. S_2	Arith. Ach. S_3	Probability I-A	Probability I-B	Probability II-A	Probability II-B	Probability III
1	110	88	99	4.8	5.9	6.8	6	3	2	-	6
2	102	94	98	5.0	6.6	5.9	5	-	2	2	4
3	100	106	103	4.6	4.6	3.6	5	2	3	-	4
4	107	88	98	4.1	3.1	6.1	3	1	2	-	8
5	113	89	101	7.1	7.6	9.1	6	2	6	2	8
6	92	97	95	4.6	5.9	4.2	2	2	2	-	2
7	113	89	101	7.9	8.2	8.0	6	5	5	2	9
8	83	95	89	4.1	4.6	3.6	6	5	2	1	6
9	110	88	99	4.6	6.1	4.0	4	6	1	2	7
10	104	90	97	5.2	6.5	5.9	5	3	4	1	3
11	84	94	89	5.0	5.4	8.0	2	6	2	-	6
12	102	93	98	3.6	4.0	4.0	3	2	1	1	1
13	110	95	103	4.4	4.6	5.4	6	5	4	1	5
14	84	89	87	4.8	4.6	4.4	4	2	1	-	3
15	117	88	103	3.6	6.1	4.2	5	4	6	2	7
16	109	93	101	5.6	5.4	7.1	6	5	3	1	4
17	85	98	92	4.6	5.9	5.7	4	1	1	-	5
18	104	86	95	4.1	3.6	4.0	1	1	3	-	3
19	109	90	100	4.4	3.6	4.9	2	5	-	-	3
20	107	88	98	6.5	6.8	7.7	5	2	6	1	8
21	96	98	97	6.3	6.3	5.7	6	3	4	1	3
22	118	83	101	4.8	5.2	4.9	3	2	2	1	5

Group 114

Subject	I.Q. L	I.Q. NL	I.Q. T	Arith. Ach. S_1	Arith. Ach. S_2	Arith. Ach. S_3	Probability I-A	Probability I-B	Probability II-A	Probability II-B	Probability III
1	91	101	95	5.8	7.2	6.3	3	-	2	-	1
2	85	97	89	5.1	6.9	6.3	5	-	3	2	2
3	95	111	102	6.4	6.0	6.7	6	6	3	2	9
4	89	87	89	5.1	4.4	4.0	3	-	1	-	5
5	96	95	95	5.8	5.4	5.3	5	1	2	-	1
6	94	80	87	4.8	5.7	7.2	-	-	2	-	4
7	108	91	101	4.5	4.8	5.8	5	2	-	-	4
8	105	80	94	5.4	5.4	5.8	2	-	3	3	8
9	85	97	90	4.5	5.1	4.9	6	5	3	1	4
10	108	90	99	4.2	4.8	4.4	2	-	4	2	3
11	102	87	95	6.4	7.2	6.7	4	5	4	2	5
12	101	101	101	6.0	6.3	5.8	1	4	1	-	2
13	99	106	103	3.9	5.1	5.3	5	5	5	-	5
14	90	102	95	6.8	5.1	4.4	4	4	6	1	1
15	87	111	97	4.8	5.4	6.7	4	6	4	-	3
16	100	105	103	4.5	5.1	7.9	4	1	3	-	3
17	85	101	91	4.5	4.4	5.3	5	-	2	-	1
18	111	90	102	6.4	5.7	6.3	6	3	6	-	9
19	107	96	103	5.1	6.3	5.3	4	5	2	-	2
20	94	89	90	3.6	7.6	4.9	6	2	5	-	8
21	96	96	95	4.5	5.4	6.3	3	-	1	-	3
22	112	75	95	4.8	6.6	6.7	6	1	2	-	5

Group 121

Subject	I.Q. L	I.Q. NL	I.Q. T	Arith. Ach. S_1	Arith. Ach. S_2	Arith. Ach. S_3	Probability I-A	Probability I-B	Probability II-A	Probability II-B	Probability III
1	100	100	100	3.5	3.3	3.8	6	1	3	-	3
2	100	98	99	3.1	3.6	2.9	2	-	1	1	3
3	97	94	96	3.6	2.3	2.9	1	-	1	1	2
4	97	87	93	2.9	2.9	3.8	-	-	-	-	1
5	99	99	98	2.5	3.9	3.8	5	-	1	-	5
6	93	94	93	3.5	3.9	4.2	2	4	2	1	2
7	105	98	102	3.7	3.9	3.8	4	-	1	-	1
8	97	102	99	3.8	4.3	4.1	2	-	1	-	3
9	83	93	86	3.5	2.1	2.9	2	-	-	1	5
10	98	100	98	3.3	2.5	3.4	1	1	2	1	2
11	107	96	101	3.1	3.9	4.9	6	1	-	2	5
12	99	104	101	3.7	2.5	3.9	1	-	-	-	3
13	99	99	99	3.7	3.0	4.6	6	2	2	-	2
14	98	107	102	3.7	4.3	4.4	6	1	1	-	3
15	102	105	103	3.6	4.5	4.0	2	2	2	2	2
16	83	99	91	3.7	2.9	4.0	2	0	-	-	-
17	95	87	93	3.6	3.6	3.8	2	1	2	-	5
18	84	92	86	3.7	3.6	3.4	3	1	-	-	5
19	96	102	99	3.7	3.9	4.1	5	1	4	-	2
20	95	94	94	2.9	2.9	3.6	3	-	1	1	-
21	93	101	95	3.3	3.3	3.8	1	-	-	-	2
22	99	110	103	5.0	4.5	4.4	5	4	2	2	3

Group 122

Subject	I.Q. L	I.Q. NL	I.Q. T	Arith. Ach. S_1	Arith. Ach. S_2	Arith. Ach. S_3	Probability I-A	Probability I-B	Probability II-A	Probability II-B	Probability III
1	99	106	103	4.9	5.0	5.5	4	2	5	-	7
2	95	99	96	4.1	2.5	3.2	1	1	3	-	5
3	93	93	93	3.7	3.9	4.0	4	1	-	-	4
4	99	95	97	5.2	3.6	3.9	4	1	-	-	1
5	105	99	101	3.4	4.5	4.1	3	-	1	-	4
6	96	97	96	3.3	4.5	3.2	3	-	1	-	5
7	112	95	103	3.5	2.7	3.9	4	3	2	2	3
8	97	102	99	3.3	3.6	4.6	4	3	2	-	5
9	95	110	100	4.3	4.3	4.6	5	-	2	-	3
10	80	95	86	2.9	2.7	3.8	1	1	-	-	3
11	97	102	99	5.4	5.4	4.7	5	2	2	-	2
12	99	106	101	4.4	5.7	5.1	4	1	1	1	4
13	97	105	100	4.6	3.9	4.4	2	2	-	2	-
14	95	98	96	3.6	3.0	3.6	4	3	1	-	1
15	102	97	99	4.9	5.4	6.1	6	3	-	-	5
16	99	105	102	4.9	4.6	5.1	4	3	1	0	5
17	93	103	96	4.1	4.1	3.9	4	-	1	1	4
18	100	105	102	3.6	3.6	3.9	3	3	1	1	2
19	94	107	98	4.1	4.6	4.2	6	1	2	1	-
20	98	108	103	6.0	5.5	4.7	6	2	1	-	1
21	92	101	95	4.8	2.9	4.1	5	1	-	-	5
22	98	98	98	3.6	5.7	4.9	2	-	3	1	4

Group 123

Subject	I.Q. L	NL	T	Arith. Ach. S_1	S_2	S_3	Probability I-A	I-B	II-A	II-B	III
1	113	87	100	4.6	4.9	3.6	2	2	3	1	6
2	106	88	97	5.4	5.4	5.1	3	5	-	-	3
3	105	83	94	4.4	4.3	3.6	3	2	2	-	5
4	107	96	102	7.7	6.3	5.9	2	2	5	1	5
5	97	77	87	5.4	5.2	3.6	3	2	2	2	5
6	93	85	89	5.4	4.6	5.4	4	-	2	-	4
7	100	96	98	6.0	6.8	7.4	6	2	1	1	6
8	117	83	100	3.8	6.8	6.1	5	2	3	1	1
9	106	97	102	5.9	7.8	6.8	5	6	2	1	7
10	103	93	98	5.9	6.1	5.9	6	4	6	1	7
11	110	90	100	4.6	4.6	4.4	4	5	1	-	6
12	110	94	102	4.1	5.2	4.6	5	1	2	-	-
13	94	95	95	4.8	4.0	3.4	5	6	1	-	4
14	109	91	100	4.8	5.9	5.4	4	5	5	1	2
15	102	88	95	2.9	4.3	4.4	2	-	2	-	2
16	98	78	88	3.3	3.6	4.2	3	1	1	2	3
17	77	93	85	5.2	5.4	3.8	4	1	3	-	5
18	102	89	96	4.4	5.2	4.1	2	-	2	-	5
19	95	92	94	2.6	4.0	4.2	3	-	3	-	3
20	107	101	103	6.5	6.1	4.9	6	3	6	2	8
21	100	102	101	4.8	7.6	7.1	4	5	-	1	4
22	100	104	102	6.3	4.0	5.6	6	5	3	1	6

Group 124

Subject	I.Q. L	I.Q. NL	I.Q. T	Arith. Ach. S_1	Arith. Ach. S_2	Arith. Ach. S_3	Probability I-A	Probability I-B	Probability II-A	Probability II-B	Probability III
1	105	95	101	7.6	6.3	7.2	5	2	3	1	-
2	96	109	102	5.1	4.8	6.7	4	3	2	1	2
3	97	92	95	5.6	4.8	7.2	2	-	2	2	5
4	94	86	88	5.4	5.7	4.4	-	-	1	3	2
5	106	94	100	4.5	3.6	5.3	5	5	4	1	1
6	96	108	101	6.6	8.5	7.4	6	5	3	2	5
7	102	83	93	4.2	5.4	4.9	5	2	4	-	2
8	106	96	101	4.5	6.3	6.3	6	5	4	1	1
9	99	107	102	7.2	6.0	4.9	4	3	3	2	8
10	101	90	96	5.8	5.1	6.7	4	5	-	-	5
11	101	102	102	4.8	5.1	4.9	3	2	6	-	5
12	92	77	84	4.8	5.4	4.9	5	3	1	-	5
13	98	97	98	5.1	5.7	4.4	4	6	2	2	1
14	102	91	97	6.6	5.4	4.9	6	4	4	1	6
15	104	101	103	4.8	6.6	6.3	5	2	3	1	4
16	94	93	93	6.4	4.8	4.4	4	2	4	1	5
17	92	111	101	4.2	6.3	4.0	4	3	3	1	1
18	97	88	92	8.0	7.8	6.3	3	-	4	-	2
19	82	111	95	6.4	5.4	8.5	3	3	2	1	2
20	88	91	88	3.6	5.1	6.3	3	5	4	1	8
21	102	94	98	6.6	8.0	7.2	3	3	3	1	4
22	106	97	102	6.0	5.1	7.9	4	4	4	1	5

Group 211

Subject	I.Q. L	I.Q. NL	I.Q. T	Arith. Ach. S_1	Arith. Ach. S_2	Arith. Ach. S_3	Probability I-A	Probability I-B	Probability II-A	Probability II-B	Probability III
1	107	101	105	2.9	2.7	3.4	2	-	2	1	-
2	107	110	108	3.6	4.3	4.6	3	-	4	-	5
3	111	106	109	3.3	4.6	4.0	4	3	1	1	1
4	108	99	105	3.1	4.1	4.0	5	1	-	1	3
5	114	107	111	2.9	2.7	3.2	1	-	1	-	5
6	108	110	110	4.6	4.5	4.4	2	-	2	-	6
7	112	105	109	5.0	5.4	7.2	6	4	4	-	3
8	101	112	106	2.9	2.9	2.9	2	-	2	-	2
9	111	102	106	3.3	3.6	2.7	1	2	1	-	2
10	116	106	112	4.3	4.3	4.4	6	2	3	2	4
11	115	104	112	3.7	2.5	2.9	3	-	1	1	1
12	124	98	111	2.2	2.7	3.6	5	3	3	1	6
13	109	106	109	3.3	4.8	5.8	2	-	4	-	5
14	101	122	110	4.9	4.8	3.6	5	1	4	-	5
15	107	107	109	3.3	3.3	4.6	4	-	-	-	5
16	109	107	109	4.4	4.6	3.8	-	3	1	1	1
17	110	101	105	4.0	4.5	3.4	1	1	2	1	2
18	119	103	112	3.6	3.6	3.2	3	-	1	2	1
19	96	127	108	3.6	5.7	6.9	6	6	3	-	8
20	111	109	113	4.4	5.1	4.6	6	-	1	-	6
21	97	118	106	2.7	4.1	6.1	4	1	2	2	5
22	106	108	108	3.5	4.3	3.8	5	1	3	-	5

Group 212

Subject	I.Q. L	NL	T	Arith. Ach. S_1	S_2	S_3	Probability I-A	I-B	II-A	II-B	III
1	112	108	112	4.6	6.1	4.4	4	1	2	1	5
2	113	99	106	5.0	5.7	4.9	2	-	3	-	4
3	114	105	111	4.0	3.3	4.2	3	1	2	-	10
4	107	111	109	4.6	5.9	4.9	6	4	1	-	7
5	95	121	105	4.3	4.6	4.1	3	2	2	1	5
6	112	104	108	5.7	6.8	6.1	4	2	3	2	5
7	116	95	105	4.9	5.9	4.6	6	5	4	-	5
8	109	104	107	4.9	5.0	4.2	1	1	2	2	6
9	116	107	112	4.4	5.7	5.8	6	6	6	-	3
10	115	95	105	4.1	5.7	5.8	6	6	-	2	9
11	113	104	110	4.0	5.5	3.9	1	-	1	1	3
12	107	107	108	4.6	6.1	7.5	4	-	6	-	8
13	110	107	110	4.3	6.1	7.2	6	4	4	-	7
14	116	102	109	4.3	5.5	5.8	5	5	6	3	5
15	109	105	107	3.8	5.0	4.9	3	3	-	-	3
16	105	110	109	4.1	5.2	5.5	2	1	-	-	2
17	120	102	111	4.6	5.5	4.9	4	4	-	-	1
18	98	114	106	3.8	5.4	4.0	5	-	3	-	5
19	104	115	109	5.3	5.5	6.2	4	6	2	2	6
20	112	101	107	3.6	3.9	3.4	4	-	2	-	4
21	107	119	113	5.6	9.5	6.5	6	3	2	2	3
22	112	99	106	4.4	5.0	4.2	4	2	1	-	6

Group 213

Subject	I.Q. L	I.Q. NL	I.Q. T	Arith. Ach. S_1	Arith. Ach. S_2	Arith. Ach. S_3	Probability I-A	Probability I-B	Probability II-A	Probability II-B	Probability III
1	113	110	112	5.8	8.2	9.6	6	4	5	3	10
2	116	100	108	6.0	7.0	6.1	5	2	5	1	3
3	110	109	110	6.3	7.3	8.0	6	5	6	2	7
4	130	81	106	6.3	7.0	6.6	4	6	6	1	5
5	105	107	106	3.8	5.4	5.9	5	2	2	-	3
6	102	109	106	5.0	5.6	5.9	5	2	1	-	3
7	126	83	106	6.8	8.2	11.1	6	6	6	5	8
8	115	107	111	4.1	5.4	3.1	1	-	1	-	4
9	108	112	110	4.6	6.3	5.1	5	6	4	1	6
10	112	101	107	5.6	5.6	5.9	4	5	2	-	5
11	132	93	113	6.2	8.8	6.8	6	5	6	2	5
12	118	91	105	5.9	7.6	6.3	6	6	3	1	7
13	111	109	110	5.0	5.9	6.1	5	2	1	-	3
14	111	108	110	6.2	7.6	9.6	4	6	3	-	5
15	128	89	109	2.9	6.8	6.6	6	6	6	1	1
16	117	101	109	6.5	6.1	4.9	4	5	-	-	8
17	116	107	112	5.4	6.8	10.1	5	6	1	1	5
18	110	103	107	3.6	5.4	4.4	4	2	-	-	2
19	124	96	110	3.3	6.6	6.6	6	5	5	3	9
20	122	91	107	5.4	8.0	9.6	4	3	5	2	5
21	124	96	110	5.2	7.6	7.4	4	2	2	1	2
22	123	101	112	4.6	7.3	5.9	6	1	2	1	4

Group 214

Subject	I.Q. L	I.Q. NL	I.Q. T	Arith. Ach. S_1	Arith. Ach. S_2	Arith. Ach. S_3	Probability I-A	Probability I-B	Probability II-A	Probability II-B	Probability III
1	101	112	106	7.8	6.9	6.7	6	2	4	2	8
2	113	105	109	8.6	7.6	7.4	6	2	5	-	7
3	108	106	107	6.0	6.3	7.2	2	2	1	1	6
4	97	122	109	5.6	5.4	8.5	4	6	4	1	8
5	115	100	100	5.1	6.0	5.8	5	3	3	-	8
6	113	101	109	5.8	6.6	7.9	6	3	6	-	5
7	106	113	109	6.4	6.3	6.7	5	3	4	-	9
8	99	119	109	4.2	8.0	5.3	6	5	-	-	8
9	115	104	111	5.4	6.0	7.2	6	6	5	1	9
10	113	106	111	6.2	5.7	7.2	6	-	5	-	2
11	118	98	110	8.9	9.9	10.4	6	6	6	5	10
12	114	106	111	6.6	8.0	10.8	6	5	3	-	9
13	118	100	100	7.6	6.6	10.4	6	6	5	1	3
14	95	116	105	3.9	4.8	4.0	5	4	-	2	10
15	106	105	106	5.4	8.0	8.5	6	4	1	0	5
16	111	112	112	6.2	7.8	7.2	6	6	2	2	4
17	115	97	108	6.8	4.8	8.2	5	5	1	-	8
18	109	101	105	6.0	6.3	9.1	6	5	5	-	6
19	94	117	105	5.6	5.7	6.3	2	6	2	1	4
20	116	100	108	5.8	6.6	9.8	6	5	6	-	7
21	104	112	108	3.6	4.8	4.4	5	5	4	1	8
22	102	118	109	5.8	6.0	5.8	5	6	5	-	7

Group 221

Subject	I.Q. L	I.Q. NL	I.Q. T	Arith. Ach. S_1	Arith. Ach. S_2	Arith. Ach. S_3	Probability I-A	Probability I-B	Probability II-A	Probability II-B	Probability III
1	113	109	113	3.1	2.9	3.8	4	3	3	-	4
2	112	99	106	3.3	2.3	3.2	4	2	4	-	5
3	113	102	109	2.9	4.5	3.4	3	2	1	-	2
4	101	111	105	2.7	2.7	4.0	3	-	-	-	1
5	116	102	109	4.0	4.6	5.5	6	1	5	1	4
6	111	99	105	3.7	5.4	4.9	4	2	1	-	2
7	103	108	105	4.4	4.1	4.0	1	-	2	1	1
8	107	101	105	3.1	3.6	4.1	5	3	-	1	3
9	104	114	110	4.1	4.1	4.0	2	4	3	-	1
10	108	114	113	4.8	3.6	5.5	6	3	6	-	1
11	115	101	109	4.3	4.8	4.7	3	4	2	2	4
12	111	111	112	3.7	5.0	4.4	6	4	2	1	6
13	118	98	108	4.0	3.0	3.9	4	1	1	-	1
14	109	106	109	4.0	4.6	4.9	6	1	3	-	1
15	110	107	110	3.7	5.7	4.4	3	5	3	-	4
16	112	111	112	2.9	2.7	3.9	4	3	2	1	1
17	113	105	110	3.1	3.9	6.1	6	-	2	2	5
18	107	104	106	3.7	1.9	2.5	3	-	1	-	4
19	107	115	112	4.0	4.3	4.9	6	2	3	2	1
20	112	105	109	3.8	3.9	3.4	1	2	1	-	3
21	102	118	109	4.4	5.0	5.1	6	5	-	2	5
22	107	115	112	3.1	3.3	3.6	3	1	3	1	1

Group 222

Subject	I.Q. L	NL	T	Arith. Ach. S_1	S_2	S_3	Probability I-A	I-B	II-A	II-B	III
1	111	107	110	5.2	5.9	5.8	6	6	4	1	4
2	111	110	113	6.2	5.0	5.8	4	4	6	1	7
3	119	104	112	4.5	6.8	5.5	2	1	1	-	2
4	102	116	109	5.0	6.5	6.2	6	6	1	1	5
5	113	104	110	4.4	6.1	4.9	4	3	3	-	4
6	110	110	111	3.6	4.3	4.6	5	4	2	1	4
7	112	107	110	4.4	4.8	4.9	5	3	1	1	4
8	115	109	113	5.6	6.3	6.9	6	6	2	2	5
9	103	109	106	3.7	5.4	5.5	6	2	1	-	4
10	107	103	105	6.4	5.5	6.5	2	2	3	1	3
11	107	100	105	4.4	3.6	5.5	-	2	3	-	3
12	101	107	105	3.8	4.5	3.8	3	3	2	-	5
13	111	113	113	2.5	4.6	4.1	5	4	3	2	3
14	108	104	107	4.0	5.4	4.1	4	1	-	-	2
15	107	113	111	5.8	6.3	7.2	4	3	4	-	4
16	101	107	105	3.3	5.9	5.8	4	3	4	-	7
17	107	104	106	5.3	6.5	7.2	5	-	3	-	3
18	112	103	108	3.6	6.8	5.5	5	4	3	2	5
19	105	111	108	4.5	3.6	4.1	3	5	4	-	2
20	117	101	110	4.8	8.0	6.9	5	6	3	1	7
21	121	96	108	4.9	4.8	5.3	4	3	5	-	7
22	96	117	105	4.1	5.2	4.1	-	-	2	-	4

Group 223

Subject	I.Q. L	I.Q. NL	I.Q. T	Arith. Ach. S_1	Arith. Ach. S_2	Arith. Ach. S_3	Probability I-A	Probability I-B	Probability II-A	Probability II-B	Probability III
1	131	87	109	5.4	5.4	5.1	5	5	5	2	5
2	106	110	108	4.6	7.3	5.7	6	4	4	4	7
3	114	105	110	6.8	7.6	7.4	6	5	3	-	5
4	117	99	108	5.8	6.8	6.6	4	5	4	-	7
5	127	97	112	5.6	7.3	6.1	6	4	4	2	4
6	106	112	109	5.4	5.9	6.1	6	4	5	1	5
7	90	128	109	8.4	7.0	8.3	6	6	4	1	7
8	108	107	108	6.5	8.2	8.0	6	5	6	2	6
9	122	100	111	6.5	7.8	8.0	4	6	2	1	7
10	123	87	105	4.6	4.9	4.4	4	4	3	1	7
11	130	91	111	4.6	3.6	4.2	2	2	2	1	6
12	124	92	108	7.4	7.0	7.1	2	3	2	2	8
13	119	94	107	6.3	6.8	6.3	6	3	5	-	6
14	123	91	107	5.8	6.6	4.9	3	-	6	-	4
15	136	90	113	6.3	5.2	6.3	6	6	6	2	9
16	108	105	105	5.4	4.4	5.1	5	2	4	-	3
17	115	97	106	6.2	8.3	6.6	6	5	6	2	9
18	106	104	105	5.6	3.6	4.9	6	1	2	1	3
19	122	89	106	5.9	6.5	5.7	6	5	4	1	4
20	109	108	109	6.6	7.6	6.3	6	6	5	2	2
21	108	99	105	4.6	4.0	4.6	4	4	2	1	2
22	117	98	108	6.0	7.6	8.0	4	5	5	-	7

Group 224

Subject	I.Q. L	I.Q. NL	I.Q. T	Arith. Ach. S_1	Arith. Ach. S_2	Arith. Ach. S_3	Probability I-A	Probability I-B	Probability II-A	Probability II-B	Probability III
1	106	103	105	6.2	6.3	7.9	5	5	1	2	6
2	119	102	112	8.0	6.6	8.2	6	4	6	1	6
3	99	122	110	6.4	5.7	6.3	6	5	6	1	8
4	121	92	109	8.9	10.3	10.8	6	6	6	5	10
5	107	108	108	4.2	4.8	5.8	6	4	5	-	2
6	102	114	107	5.4	4.8	7.4	4	-	2	2	5
7	109	103	107	9.2	8.5	8.2	6	6	5	1	10
8	116	108	113	5.6	6.3	7.4	4	6	5	2	6
9	97	113	105	6.2	5.1	5.3	5	3	2	2	7
10	113	108	111	6.0	6.0	5.7	5	4	4	-	6
11	108	106	108	5.6	5.7	5.8	5	3	1	1	4
12	120	103	113	8.9	7.6	8.5	6	5	5	2	10
13	114	93	105	6.0	5.1	8.2	6	3	6	1	10
14	125	95	112	6.6	6.3	7.4	6	4	6	1	8
15	97	121	108	7.2	7.8	9.1	6	6	5	-	6
16	123	97	113	4.8	7.2	7.2	2	5	2	-	4
17	119	100	112	7.6	6.9	8.2	5	4	6	1	2
18	111	115	113	6.6	9.6	9.8	6	5	4	3	5
19	104	117	111	5.8	8.2	7.9	6	5	2	-	1
20	108	106	108	5.4	5.7	7.4	5	4	3	-	8
21	105	119	112	5.1	6.3	5.8	4	1	6	1	2
22	105	113	110	8.2	5.4	8.2	3	3	2	-	4

Group 311

Subject	I.Q. L	I.Q. NL	I.Q. T	Arith. Ach. S_1	Arith. Ach. S_2	Arith. Ach. S_3	Probability I-A	Probability I-B	Probability II-A	Probability II-B	Probability III
1	124	128	130	3.8	4.6	4.9	1	-	3	1	2
2	129	114	125	3.5	5.4	5.8	5	6	5	-	8
3	114	112	114	2.7	3.3	4.4	2	5	4	-	3
4	113	115	116	4.1	5.5	4.7	6	1	22	2	3
5	112	111	114	4.0	3.6	5.3	4	2	2	1	3
6	129	121	123	4.0	4.6	4.6	6	4	5	2	6
7	115	120	120	4.6	7.6	7.6	6	5	6	2	8
8	126	108	120	3.7	5.4	6.1	6	6	4	1	5
9	121	121	125	4.8	4.3	5.1	3	-	3	-	7
10	125	119	125	3.5	6.8	4.9	6	6	4	1	3
11	113	131	123	4.3	6.1	7.2	3	4	6	1	9
12	111	124	120	4.3	5.9	4.9	5	-	1	-	3
13	109	121	116	4.5	5.2	4.7	2	-	4	-	4
14	133	123	131	3.8	6.3	5.3	4	5	5	-	10
15	121	109	117	2.9	3.9	3.4	4	3	1	-	1
16	121	123	126	3.8	5.5	4.7	4	5	4	3	8
17	117	116	119	3.5	5.2	6.5	5	2	5	1	7
18	125	107	117	4.4	5.7	6.9	6	6	6	2	7
19	116	108	114	3.6	4.6	4.7	3	-	1	1	2
20	111	121	118	3.5	5.5	5.3	5	4	3	1	7
21	121	108	118	1.6	2.9	4.0	2	4	2	2	1
22	117	113	117	3.8	5.7	6.1	4	2	6	-	5

Group 312

Subject	I.Q. L	NL	T	Arith. Ach. S_1	S_2	S_3	Probability I-A	I-B	II-A	II-B	III
1	120	107	114	5.2	5.2	4.9	4	3	4	-	6
2	117	117	119	4.4	6.3	5.8	4	3	6	2	9
3	107	127	118	6.0	8.5	6.1	4	5	5	-	5
4	129	121	129	4.8	6.3	6.1	4	6	6	-	4
5	112	123	119	8.6	7.6	9.5	4	4	5	-	7
6	110	125	119	4.5	7.1	5.3	6	4	2	1	8
7	127	106	118	3.1	6.3	5.1	4	4	3	-	5
8	116	134	128	7.7	9.5	9.5	5	6	6	6	10
9	128	115	124	5.3	7.6	9.5	5	5	5	-	7
10	121	110	117	4.3	6.8	6.5	6	6	3	3	8
11	113	112	115	3.7	6.3	4.2	4	-	4	-	4
12	121	135	130	4.8	7.6	9.0	6	6	6	1	9
13	116	125	123	5.2	9.5	8.0	6	5	4	2	7
14	113	112	115	3.3	3.6	4.1	4	4	3	-	5
15	127	122	131	4.3	9.5	5.8	6	6	6	5	10
16	129	121	129	3.5	8.0	5.1	4	6	5	1	8
17	120	107	117	4.6	5.5	4.6	6	3	4	-	2
18	114	122	119	4.1	3.9	3.6	3	4	2	-	6
19	121	121	125	5.7	6.1	6.5	6	6	5	-	5
20	109	119	116	4.5	5.7	4.6	4	5	5	1	9
21	109	119	114	3.8	5.9	5.8	4	-	2	-	2
22	116	108	114	3.6	6.1	4.9	3	4	4	1	6

Group 313

Subject	I.Q. L	I.Q. NL	I.Q. T	Arith. Ach. S_1	Arith. Ach. S_2	Arith. Ach. S_3	Probability I-A	Probability I-B	Probability II-A	Probability II-B	Probability III
1	117	121	117	4.6	4.6	6.5	6	2	3	-	6
2	142	118	130	6.8	7.3	7.7	2	5	3	1	7
3	118	116	117	4.6	6.5	6.6	2	2	4	-	8
4	110	134	122	7.1	8.5	8.6	4	3	4	-	5
5	135	109	122	6.6	8.0	8.3	6	5	5	2	9
6	139	99	119	5.8	8.0	9.1	6	5	5	-	1
7	130	107	119	6.0	6.5	6.6	6	5	4	2	8
8	122	118	120	5.6	6.8	8.0	6	6	5	1	7
9	122	118	120	4.8	5.4	5.4	4	5	4	2	2
10	143	119	131	4.4	6.3	8.0	6	5	4	2	10
11	135	134	135	7.7	8.5	9.6	6	6	6	3	7
12	134	107	121	6.0	7.8	9.1	6	5	5	-	1
13	131	130	131	5.6	6.8	6.5	6	5	5	3	10
14	112	118	115	4.1	4.6	4.2	4	3	1	-	5
15	143	98	121	7.1	7.8	8.6	4	6	3	-	5
16	129	124	127	6.3	7.8	7.7	5	6	6	1	6
17	130	107	119	4.1	6.8	6.8	3	5	5	2	2
18	144	124	134	6.3	8.8	10.1	6	4	2	1	7
19	124	116	120	6.8	6.6	9.1	6	6	5	2	8
20	138	89	114	4.8	6.5	6.3	6	3	6	1	9
21	127	134	131	7.7	7.8	8.0	6	6	6	-	9
22	145	135	140	7.4	8.2	10.6	6	5	6	2	7

Group 314

Subject	I.Q. L	I.Q. NL	I.Q. T	Arith. Ach. S_1	Arith. Ach. S_2	Arith. Ach. S_3	Probability I-A	Probability I-B	Probability II-A	Probability II-B	Probability III
1	132	132	134	12.1	12.0	12.5	6	6	6	2	10
2	125	126	127	11.9	11.4	11.9	6	6	6	1	10
3	117	116	118	6.0	6.3	8.2	6	6	4	1	8
4	126	134	132	7.2	8.5	8.2	6	5	6	3	10
5	119	125	123	7.8	8.0	11.1	5	6	6	2	8
6	131	127	131	12.7	12.7	12.5	6	5	6	4	9
7	118	118	120	7.8	6.3	10.4	6	5	6	1	9
8	122	122	124	5.1	6.6	8.2	6	5	4	-	7
9	122	115	121	5.8	6.0	4.4	6	4	5	1	7
10	120	121	121	7.9	7.8	9.1	6	6	6	2	8
11	118	111	117	4.5	6.6	10.4	6	4	3	-	3
12	115	112	115	4.8	6.0	5.8	5	-	1	1	6
13	122	127	126	8.9	8.8	11.9	6	6	5	6	10
14	137	128	128	6.2	6.9	7.9	4	6	5	5	8
15	123	107	118	4.5	5.4	5.8	6	5	4	1	4
16	126	116	123	8.6	9.6	7.4	6	5	6	4	10
17	122	120	123	5.6	6.3	7.9	6	6	6	5	7
18	114	117	117	6.4	6.0	4.4	5	6	3	3	4
19	125	120	124	4.2	6.3	7.2	6	3	6	1	8
20	122	111	119	6.2	7.2	7.4	6	4	6	1	9
21	115	122	119	6.4	9.2	11.3	5	6	4	4	9
22	117	116	118	8.0	11.1	10.8	6	6	6	2	8

Group 321

Subject	I.Q. L	I.Q. NL	I.Q. T	Arith. Ach. S_1	Arith. Ach. S_2	Arith. Ach. S_3	Probability I-A	Probability I-B	Probability II-A	Probability II-B	Probability III
1	121	129	129	3.8	5.5	4.2	6	6	6	1	7
2	120	118	121	3.8	5.5	6.5	4	5	2	1	6
3	121	126	126	3.7	5.0	4.6	4	2	2	1	2
4	114	113	114	3.1	4.6	4.1	4	-	-	-	6
5	108	119	115	3.7	5.2	4.6	5	6	6	-	6
6	108	128	119	3.6	5.7	4.6	5	4	6	-	3
7	124	115	122	3.8	3.6	4.6	6	-	5	-	3
8	125	103	115	4.5	5.9	5.1	3	4	5	1	6
9	128	107	119	4.0	4.6	3.9	4	6	5	-	6
10	129	113	123	3.8	5.2	4.0	6	6	4	-	6
11	114	117	117	3.6	5.9	6.1	3	2	-	1	6
12	118	112	116	4.1	4.5	4.2	2	2	2	1	3
13	105	121	114	2.5	5.2	4.6	3	2	4	-	3
14	128	120	128	4.4	5.4	4.2	5	1	4	-	6
15	113	126	122	4.0	5.0	4.4	3	1	6	1	6
16	121	107	116	3.1	5.7	4.7	6	1	5	-	1
17	113	125	121	4.8	5.5	5.5	5	3	3	2	3
18	120	109	119	4.0	4.1	4.2	3	2	-	2	2
19	111	117	117	3.5	5.0	4.9	5	2	2	-	1
20	111	118	116	3.6	4.5	5.8	6	5	-	-	5
21	120	121	123	3.1	5.7	5.8	5	6	3	2	4
22	121	111	120	4.1	5.5	5.8	6	5	6	1	6

Group 322

Subject	I.Q. L	I.Q. NL	I.Q. T	Arith. Ach. S_1	Arith. Ach. S_2	Arith. Ach. S_3	Probability I-A	Probability I-B	Probability II-A	Probability II-B	Probability III
1	111	120	118	4.8	4.1	4.2	6	5	4	-	6
2	121	117	121	4.3	5.2	4.2	4	5	4	-	4
3	114	116	118	5.0	5.4	5.3	5	6	4	-	3
4	116	111	116	5.4	6.5	5.8	5	3	-	-	8
5	121	104	114	5.6	5.2	5.3	6	3	5	-	6
6	116	130	126	5.3	9.5	7.5	6	4	2	2	9
7	114	128	124	5.7	6.5	8.5	4	6	6	1	10
8	107	120	115	5.2	8.0	6.9	6	6	4	-	5
9	119	111	118	4.0	5.4	5.8	6	6	-	1	5
10	122	131	129	6.2	7.6	8.5	6	5	3	-	8
11	126	111	121	5.0	5.4	5.3	6	5	1	1	4
12	123	117	123	5.8	8.5	7.5	6	6	6	2	5
13	127	122	127	6.4	7.6	8.5	6	6	5	2	8
14	126	126	130	9.5	6.8	8.5	6	6	6	4	6
15	117	121	121	4.1	5.9	3.8	6	6	2	1	6
16	115	117	119	4.8	6.5	6.9	6	6	1	-	7
17	108	124	116	3.6	5.7	4.9	4	-	2	-	2
18	109	123	116	8.2	7.1	8.0	6	6	6	-	7
19	122	107	115	5.6	5.5	5.1	5	1	5	-	9
20	121	104	114	4.6	6.1	4.6	2	5	5	1	5
21	116	108	114	5.2	8.0	6.9	6	3	4	1	2
22	125	114	123	5.6	5.5	7.5	4	6	3	-	6

Group 323

Subject	I.Q. L	I.Q. NL	I.Q. T	Arith. Ach. S_1	Arith. Ach. S_2	Arith. Ach. S_3	Probability I-A	Probability I-B	Probability II-A	Probability II-B	Probability III
1	98	132	115	5.2	5.9	4.6	3	4	3	-	4
2	117	114	116	6.3	6.6	6.3	6	1	6	1	4
3	118	116	117	6.6	8.5	6.5	6	6	4	-	9
4	122	107	115	5.4	6.6	4.4	6	4	2	1	7
5	118	124	121	5.6	6.3	6.1	6	5	3	2	2
6	116	122	119	6.0	5.6	6.1	6	6	3	-	8
7	134	106	120	5.4	6.8	7.1	4	6	5	1	8
8	135	97	116	7.4	7.8	8.6	6	5	6	5	7
9	143	105	124	7.4	8.5	9.6	6	6	5	-	5
10	139	134	137	6.2	9.5	8.6	6	5	3	3	9
11	117	131	124	4.1	4.6	4.0	6	-	-	-	6
12	135	124	130	5.4	7.8	8.6	6	4	6	-	7
13	130	133	132	4.6	6.3	7.1	6	6	5	-	6
14	124	139	132	6.5	10.3	8.0	6	3	6	1	9
15	141	135	138	6.8	9.5	8.0	4	6	6	-	9
16	119	121	120	5.2	6.8	7.1	4	4	6	-	4
17	133	128	131	6.0	8.0	9.6	6	6	6	3	9
18	124	127	126	6.0	6.6	6.3	5	5	2	1	9
19	120	108	114	6.8	7.3	5.6	6	3	4	-	9
20	113	130	122	5.0	5.9	6.3	6	2	2	-	5
21	121	132	127	7.4	6.8	7.4	6	5	6	1	6
22	130	108	119	5.9	6.3	8.3	4	3	4	-	2

Group 324

Subject	I.Q.			Arith. Ach.			Probability				
	L	NL	T	S_1	S_2	S_3	I-A	I-B	II-A	II-B	III
1	118	117	118	6.8	6.6	7.9	6	5	5	-	8
2	119	117	119	7.8	6.6	7.4	6	4	-	-	8
3	129	98	115	6.6	8.2	7.4	6	5	5	1	7
4	118	124	121	8.9	9.2	9.8	5	6	6	1	8
5	115	114	116	6.0	7.6	7.9	6	6	6	-	8
6	115	128	122	10.0	10.7	11.1	5	6	6	-	9
7	119	112	117	6.8	7.8	7.2	6	5	6	-	5
8	130	121	128	6.2	7.8	7.4	6	6	6	1	9
9	117	116	118	6.8	7.6	6.7	5	5	3	-	5
10	129	114	125	6.6	8.0	7.4	6	6	6	-	9
11	118	120	120	6.2	6.6	8.2	6	6	6	2	8
12	116	116	118	7.8	6.6	6.7	6	6	5	1	4
13	128	123	128	8.2	7.8	11.3	6	6	6	3	9
14	129	125	129	7.8	6.9	9.1	6	4	6	1	9
15	123	124	125	7.2	7.6	7.4	6	5	6	-	2
16	120	119	122	8.2	9.2	8.2	6	5	6	1	8
17	124	125	126	8.6	7.8	10.4	6	6	5	2	6
18	124	126	127	4.8	7.6	7.2	6	3	6	2	9
19	115	119	118	5.1	8.5	8.2	6	5	5	3	9
20	118	117	119	7.6	5.7	8.5	6	5	6	2	8
21	123	124	125	9.6	9.9	10.4	6	5	5	3	7
22	121	116	121	6.2	7.2	10.8	5	4	5	6	8

BIBLIOGRAPHY

Adler, M. J. Some Educational Implications of the Theories of Jean Piaget and J. S. Bruner. Canadian Educational Research Digest, 1965, 5: 5-13.

Baker, Frank. An Intersection of Test Score Interpretation and Item Analyis. Journal of Educational Measurement, 1964, 1: 23-28.

Baker, Frank. Empirical Determination of Sampling Distributions of Item Discrimination Indices and A Reliability Coefficient. Department of Educational Psychology, School of Education, University of Wisconsin, Contract OE-2-10-071, November, 1962.

Baker, Frank. Test Analysis Package: A Program for the CDC 1604-3600 Computers. Laboratory of Experimental Design, Department of Educational Psychology, University of Wisconsin, 1966.

Beard, R. M. The Nature and Development of Concepts, I. Educational Review, 1960, 13: 157-175.

Beard, R. M. The Nature and Development of Concepts, II. Educational Review, 1961, 13: 12-26.

Beard, R. M. The Order of Concept Development Studies in Two Fields. Educational Review, 1962-1963, 15: 105-117, 228-237.

Beard, R. M. Further Studies in Concept Development. Educational Review, 1964, 17: 41-58.

Bell, E. T. Men of Mathematics. New York: Simon and Schuster Company, 1937.

Bock, R. Darrell. Programming Univariate and Multivariate Analysis of Variance. Technometrics, 1963, 5: 95-117.

Cambridge Conference on School Mathematics. Goals for School Mathematics. Boston: Houghton Mifflin Company, 1963.

Cambridge Conference on School Mathematics. "Reports of the 1965 Summer Study." Unpublished mimeograph report, May, 1966, Section II-Chapter 4.

Carpenter, T. E. A Pilot Study for a Quantitative Investigation of Jean Piaget's Original Work on Concept Formation. Educational Review, 1955, 7: 142-148.

Carnap, Rudolph. What is Probability? Scientific American, 1953, 189: 128-138.

Cochran, William G. Analysis of Covariance: Its Nature and Uses. Biometrics, 1957, 3: 261-281.

Cohen, John and Hansel, C. E. M. Risk and Gambling. London: Longmans, Green, 1956.

Cohen, John. Subjective Probability. Scientific American, 1957, 197: 128-138.

Cohen, John, Dearnaley, E. J. and Hansel, C. E. M. Measures of Subjective Probability. British Journal of Psychology, 1957, 48: 271-275.

Cohen, John. Chance, Skill and Luck: the Psychology of Guessing and Gambling. Baltimore: Penguin Books, 1960.

Commission on Mathematics College Entrance Examination Board. Program for College Preparatory Mathematics. New York: College Entrance Examination Board, 1959.

Commission on Mathematics College Entrance Examination Board. Introductory Probability and Statistical Inference, An Experimental Course. Princeton: Educational Testing Service, 1959.

Davies, C. M. Development of the Probability Concept in Children. Child Development, 1965, 36: 779-788.

Deans, Edwina. Elementary School Mathematics: New Directions. Washington, D. C.: U. S. Office of Education, 1963.

Donaldson, M. Review of Piaget and Inhelder. British Journal of Psychology, 1960, 51: 181-184.

Duncan, Ernest R., Capps, Lelon R., Doliciani, Mary P., Quast, W. G. and Zweng, Marilyn. Modern School Mathematics Structure and Use, Books 5 and 6. Boston: Houghton Mifflin Company, 1967.

Fehr, Howard, Bunt, L. and Grossman, G. An Introduction to Sets, Probability and Hypothesis Testing. Boston: Heath Book Company, 1964.

Ferguson, George A. Statistical Analysis in Psychology and Education. New York: McGraw-Hill Book Company, Inc., 1959.

Finn, Jeremy D. Multivariance: Fortran Program for Univariate and Multivariate Analysis of Variance and Covariance. Buffalo: State University of New York at Buffalo, Department of Educational Psychology, May, 1967.

Flavell, J. H. The Developmental Psychology of Jean Piaget. Princeton, New Jersey: D. Van Nostrand Company, Inc., 1963.

Goldberg, Samuel. An Introduction to Probability. Englewood Cliffs, New Jersey: Prentice Hall Book Company, 1960.

Gratch, G. The Development of the Expectation of the Nonindependence of Random Events in Children. Child Development, 1959, 30: 217-227.

Hartung, Maurice, Van Engen, Henry, Gibb, E. G., Stochl, James, Knowles, L. and Walch, R. Seeing Through Arithmetic 6. Chicago: Scott, Foresman and Company, 1968.

Jones, Lyle V. Some Illustrations of Psychological Experiments Designed for Multivariate Statistical Analysis. Report Number 28, the Psychometric Laboratory, University of North Carolina, December, 1960.

Kass, Norman. Risk and Decision Making as a Function of Age, Sex, and Probability Preference. Child Development, 1964, 35: 577-582.

Kelley, T. L., Madden, R., Gardner, E. F. and Rudman, H. C. Stanford Achievement Test. New York: Harcourt Brace and World, Inc., 1964.

Kemeny, John, Mirkil, H., Snell, J. L. and Thompson, G. L. Finite Mathematical Structures. Englewood Cliffs, New Jersey: Prentice Hall Book Company, 1959.

King, A. C. and Read, C. B. Study Through Biographies of the History of the Elementary Concepts of Probability. School Science and Mathematics, 1962, 62: 165-176.

Komorita, S. S. Factors Which Influence Subjective Probability. Journal of Experimental Psychology, 1959, 58: 386-389.

Leake, Lowell. "The Status of Three Concepts of Probability in Children of 7th, 8th and 9th Grades." Unpublished Ph.D. dissertation, University of Wisconsin, 1962.

Lovell, Kenneth. The Growth of Basic Mathematical and Scientific Concepts in Children. New York: Philosophical Library, 1961.

Lovell, Kenneth and Butterworth, J. B. Abilities Underlying the Understanding of Proportionality. Mathematics Teaching, 1966, 37.

McSwain, E. T., Brown, Kenneth E., Gundlach, Bernard H., and Cooke, Ralph J. Arithmetic 4, 5 and 6. River Forest, Illinois: Laidlow Brothers, 1965.

Messick, S. J. and Solley, C. M. Probability Learning in Children: Some Exploritory Studies. Journal of Genetic Psychology, 1957, 90: 23-32.

Mosteller, Frederick, Rourke, R. and Thomas, G. Probability: A First Course. Reading, Massachusetts: Addison-Wesley Book Company, 1961.

Newman, J. The World of Mathematics. New York: Simon and Schuster Company, 1956.

Ojemann, Ralph, Maxey, E. J. and Snider, B. C. Effects of a Program of Guided Learning Experiences in Developing Probability Concepts at 3rd Grade. Journal of Experimental Education, 1965, 33: 321-330.

Ojemann, Ralph, Maxey, E. J. and Snider, B. C. Effects of Guided Learning Experiences in Developing Probability Concepts at the Fifth Grade Level. Preceptional and Motor Skills, 1965, 415-427.

Osborn, Roger, DeVault, M. V., Boyd, C. C. and Houston, W. R. Extending Mathematical Understanding. Columbus, Ohio: Merrill Book Company, 1961.

Page, David A. Probability. The Growth of Mathematical Ideas, Grades K-12. Twenty-Fourth Yearbook of the National Council of Teachers of Mathematics. Washington, D. C.: The Council, 1959, 229-271.

Parzen, Emanuel. Modern Probability Theory and its Applications. New York: John Wiley and Sons, 1960.

Piaget, Jean and Inhelder, Barbel. La Genése de l'Idée de Hasard chez l'Enfant. Paris: Presses Universitaires de France, 1951.

Pire, G. Notion du Hasard et Developpement Intellectual. Enfrance, 1958, 131-143.

Read, C. B. Elementary Probability. School Science and Mathematics, 1963, 63: 119-132.

Ross, B. M. and Levy, N. Patterned Predictions of Chance Events by Children and Adults. Psychological Reports, 1958, 4: 87-126.

Sanderstrom, H. E. Probability in the Classroom. School Science and Mathematics, 1964, 64: 679-682.

School Mathematics Study Group. Probability for Primary Grades. Pasadena, California: A. C. Vroman, Inc., 1966

School Mathematics Study Group. Probability for Intermediate Grades. Pasadena, California: A. C. Vroman, Inc. 1966.

Smith, Rolland R. Probability in the Elementary School. Enrichment Mathematics for the Grades. Twenty-Seventh Yearbook of the National Council of Teachers of Mathematics. Washington, D. C.: The Council, 1963, 127-133.

Stevenson, H. W. and Weir, M. W. The Role of Age and Verbalization in Probability Learning. American Journal of Psychology, 1963, 76: 299-305.

Stevenson, H. W. and Zigler, E. F. Probability Learning in Children. Journal of Experimental Psychology, 1958, 56: 185-192.

Sullivan, Elizabeth T., Clark, Willis W. and Tiegs, Ernest W. California Short-Form Test of Mental Maturity, 1963 Revision. Monterey, California: California Test Bureau, 1964.

Suppes, Patrick, Kyser, Dolly and Braithwaite, Catherine. Sets and Numbers, Books 4, 5 and 6. New York: Singer Company, Inc., 1966.

Van Engen, Henry. The Formation of Concepts. The Learning of Mathematics, Its Theory and Practice. Twenty-First Yearbook of the National Council of Teachers of Mathematics. Washington, D. C.: The Council, 1953.

Van Engen, Henry, Hartung, Maurice, Trimble, Harold, Berger, Emil and Cleveland, Ray. Seeing Through Mathematics, Books 1 and 2. Chicago: Scott, Foresman and Company, 1961.

Van Engen, Henry, Hartung, Maurice, Trimble, Harold, Berger, Emil and Cleveland, Ray. Teaching Guide, Seeing Through Mathematics, Book 2. Chicago: Scott, Foresman and Company, 1963.

Weir, M. W. The Effects of Age and Instruction on Children's Probability Learning. Child Development, 1962, 33: 729-735.

Wilkinson, J. D. and Nelson, O. Probability and Statistics: Trial Teaching in 6th Grade. The Arithmetic Teacher, 1966, 13: 100-106.

Williams, H. E. The Place of Probability and Statistics in the Mathematics Curriculum of the Senior High School and Junior College. The Mathematics Teacher, 1961, 54: 316-320.

Willoughby, Stephen S. Contemporary Teaching of Secondary School Mathematics. New York: John Wiley and Sons, 1967.

Winer, B. J. Statistical Principals in Experimental Design. New York: McGraw-Hill Book Company, 1962.

Wolf, Frank L. Elements of Probability and Statistics. New York: McGraw-Hill Book Company, 1962.

Yost, Patricia A., Siegel, A. and Andrews, J. Nonverbal Probability Judgements by Young children. Child Development, 1962, 33: 769-780.